

LANGUAGECERT ACADEMIC CONCORDANCE REPORT

October 2024

Leda Lampropoulou
Michael Milanovic
LANGUAGECERT

Johnathan Jones
Anthony Green
University of Bedfordshire

Authors

Leda Lampropoulou, Research Manager, LANGUAGECER

Michael Milanovic, Chair of the Board of Directors, LANGUAGECER

Anthony Green, Director, Centre for Research in English Language Learning and Assessment (CRELLA), University of Bedfordshire

Johnathan Jones, Lecturer, Centre for Research in English Language Learning and Assessment (CRELLA), University of Bedfordshire

External Validation panel

Anthony Green, Professor in Language Assessment, Director of the Centre for Research in English Language Learning and Assessment (CRELLA), University of Bedfordshire, United Kingdom

Liyong Cheng, Professor of Language Education and Assessment, Graduate Faculty Director, Assessment and Evaluation Group, Queen's University, Canada

Sara Cushing, Professor of Applied Linguistics, ESL Senior Faculty Associate for Assessment of Student Learning, Georgia State University, USA

Jason Fan, Senior Research Fellow, Deputy Director of the Language Testing Research Centre (LTRC), University of Melbourne, Australia

James Tognolini, Professor and Director of the Centre for Educational Measurement and Assessment (CEMA), University of Sydney, Australia

Suggested citation

Lampropoulou, L., Milanovic, M., Jones, J., Green, A. (2024), *LANGUAGECER Academic Concordance Report*. LANGUAGECER.

Table of Contents

List of tables	3
List of figures	3
Foreword	4
LANGUAGECERT	6
Concordance studies overview	6
LANGUAGECERT Academic – IELTS Academic Concordance study	7
Section 1: Content Comparisons.....	10
Summary of LANGUAGECERT Academic and IELTS Academic tests	10
LANGUAGECERT Academic	10
IELTS Academic.....	10
Reporting results and feedback to participants.....	11
Content comparisons	12
Speaking comparison	12
Summary of Speaking - Key similarities and distinguishing features	16
Listening comparison	19
Summary of Listening - Key similarities and distinguishing features	23
Writing comparison	25
Summary of Writing - Key similarities and distinguishing features	28
Reading comparison	30
Summary of Reading - Key similarities and distinguishing features.....	34
Conclusion to Content Comparison Section	36
Section 2: Statistical analyses and results	37
Test taker demographics	37
Sample population	37
Age	38
Gender	38
Nationality.....	39
First language	42
Self-reported proficiency	42
Sequence and Exam interval	43
Familiarity.....	44
Test takers' reasons for taking the test.....	45
Score distribution summary statistics	45
Test taker performance analysis.....	46
Correlations	46
Comparing test scores.....	47
Skill comparison	48
Recommendations for interpretation and use of linkage results for test users	52
Population invariance.....	52
Concordance results discussion	55
Conclusion to Statistical analyses and results section	58
References	59
Appendix A: Exam interval analysis	62
Appendix B: Concordance study – Candidate Questionnaire	64

List of tables

Table 1. English language tests concordance studies with IELTS Academic	7
Table 2. Speaking test comparison of LCA IELTS (Academic and General Training)	12
Table 3. Listening test comparison of LCA and IELTS (Academic and General Training).....	19
Table 4. Writing test comparison of LCA and IELTS Academic	25
Table 5. Reading test comparison of LCA and IELTS Academic: Task Features	30
Table 6. Reading test comparison of LCA and IELTS Academic: Features of the Input Text	33
Table 7. IELTS Academic mean total scores by gender compared to published data	39
Table 8. IELTS Academic mean skill component scores by gender compared to published data	39
Table 9. Test taker Nationality	40
Table 10. IELTS Academic mean total scores by nationality compared to published data	40
Table 11. IELTS Academic mean skill component scores by nationality	41
Table 12. Test taker first language background.....	42
Table 13. Test taker familiarity with the LANGUAGECER Academic test at exam registration	45
Table 14. Test taker familiarity with the IELTS Academic test at exam registration.....	45
Table 15. Test takers' reasons for taking the LANGUAGECER Academic and the IELTS Academic tests..	45
Table 16. Distribution statistics	46
Table 17. Subskill descriptive statistics comparison.....	46
Table 18. Correlations between LANGUAGECER Academic and IELTS Academic	47
Table 19. Overall alignment table	48
Table 20. Listening skill alignment table	50
Table 21. Reading skill alignment table	50
Table 22. Writing skill alignment table.....	51
Table 23. Speaking skill alignment table	51
Table 24. Population Invariance - Overall	53
Table 25. Population Invariance - Speaking.....	53
Table 26. Population Invariance – Listening	53
Table 27. Population Invariance – Reading.....	54
Table 28. Population Invariance – Writing	54
Table 29. Duration between exams within a 90-day period.....	62
Table 30. Duration between exams > 90 days	62

List of figures

Figure 1. Overseas migrant arrivals - age and sex.....	38
Figure 2. Overall score relationship between LANGUAGECER Academic and IELTS Academic.....	47
Figure 3. IELTS Writing scores distribution	56
Figure 4. LANGUAGECER Academic Writing scores distribution.....	56
Figure 5. IELTS Listening scores distribution	57
Figure 6. LANGUAGECER Academic Listening scores distribution	57

Foreword

LANGUAGECERT's mission is to provide high-quality, accessible, and fit-for-purpose English language assessments that empower individuals and institutions worldwide. As part of this mission, it is with great pride that we present this comprehensive concordance study, a milestone in the journey of the LANGUAGECERT Academic test since its initial pilot in 2022. Over the course of two years, this study rigorously examined the relationship between LANGUAGECERT Academic and IELTS Academic, with the aim of providing stakeholders with data-driven evidence to support their academic and migration-related decisions.

The study, involving over 1,000 test takers from diverse linguistic and cultural backgrounds, was conducted with meticulous attention to research design and methodology. The research was grounded in a dual-approach methodology: a qualitative content analysis and a quantitative statistical analysis. The first phase of the study focused on a detailed comparison of the tests' constructs, task types, and scoring methods. This content analysis confirmed a high degree of similarity in the targeted constructs and language use domains, establishing a meaningful basis for comparing test taker performances.

Building on this foundation, the second phase employed a robust quantitative approach, adopting a single-group design to ensure consistency and reliability. Test takers, representing over 46 nationalities and 35 first languages, completed both tests within a three-month period, with careful counterbalancing of test order to minimize sequencing effects. Their official IELTS test report forms were collected to ensure data accuracy, and rigorous statistical techniques, including Pearson correlations and equipercentile rankings, were used to create score equivalence pairs.

The findings reveal a strong overall correlation between the two tests, reflecting their shared focus on measuring English proficiency in academic domains. In line with good practice in concordance studies (Knoch & Fan, 2024), we emphasise that score equivalencies, while based on empirical evidence, should be considered indicative. Decision-makers are encouraged to use these results in conjunction with other sources of evidence to make well-informed decisions.

This report reflects the collaborative efforts of leading experts in the field, including our internal Research Team and the Centre for Research in English Language Learning and Assessment (CRELLA) at the University of Bedfordshire. It serves as a testament to LANGUAGECERT's dedication to evidence-based practices and continuous improvement. I would like to thank the authors for their meticulous work on the design and development of the study. We are also deeply grateful to the members of our External Validation Panel (previously Concordance Studies Advisory panel), whose oversight ensured and confirmed the study's robustness and the reliability of its findings.

This concordance study underscores LANGUAGECERT Academic as a fit-for-purpose qualification

for those seeking to demonstrate their English proficiency. LANGUAGECERT remains committed to advancing the science of language testing, promoting fairness, reliability, and accessibility for all test takers. This report reflects our dedication to supporting stakeholders with rigorous research, empowering them to make informed decisions in an increasingly globalised world.

We hope this report serves as a valuable resource and contributes to the ongoing discourse on language test comparability.

Marios Molfetas

Chief Languages Officer

LANGUAGECERT

LANGUAGECERT

LANGUAGECERT's established international English qualifications are recognised by employers, educational institutions, and professional bodies worldwide. LANGUAGECERT and its test centre partner, Prometric, have a global network of 193 assured facilities for secure test delivery in 116 countries. LANGUAGECERT is authorized by UK Visas and Immigration (UKVI) to deliver Home Office approved, Secure English Language Tests (SELTs) in the UK and globally and has been granted Awarding Organisation recognition by Ofqual (a non-ministerial government department that regulates qualifications, exams and tests in England) and by its Welsh counterpart, Qualifications Wales.

LANGUAGECERT has introduced an English Language Test system for people intending to work, study or live where English is used as the primary language for communication. The system comprises LANGUAGECERT Academic, used for higher education admissions purposes, and LANGUAGECERT General, intended for work or vocational immigration purposes. LANGUAGECERT has developed a systematic plan to demonstrate that the tests are valid (i.e., fit-for-purpose), reliable, and secure. This plan includes internal on-going research analyses as well as external validation projects, such as the benchmarking of the LANGUAGECERT Global Scale against the Common European Framework of Reference for Languages (CEFR) (Council of Europe, 2001).

As part of this plan, LANGUAGECERT commissioned a study into the extent of concordance in content and performance between LANGUAGECERT Academic and IELTS Academic and between LANGUAGECERT General and IELTS General Training. Materials sourced for analysis include publicly available manuals and online documentation (<https://ielts.org/researchers/our-research/test-statistics>), research and peer reviewed articles, and test scores. This report focuses on the results from the investigation of the concordancing between LANGUAGECERT Academic and IELTS Academic.

Concordance studies overview

Concordance studies play a crucial role in language assessment and testing by investigating the comparability and alignment of different language proficiency tests. These studies aim to establish empirical evidence of the equivalence or similarity between test scores from different language assessments, enabling informed decision-making regarding test interpretation and usage. They are crucial for several reasons, including, among others, establishing score mapping, evaluating test comparability, informing test interpretation, as well as enhancing test development. Results from concordance studies can inform policymakers and educational institutions in setting standards for language proficiency and selecting appropriate tests for different educational purposes.

Over the past two decades, substantial interest has been seen in linking different tests. Linking methods may be viewed on a cline (Pommerich & Dorans, 2004). In equating, tests measure the same construct with equal reliability. Analysis is conducted through equating raw scores from both tests. Scores are reported on the same scale and can be used interchangeably. Concordancing establishes a relationship between scores on tests that are built to different specifications. They measure similar but not identical constructs. Scores from two tests linked through concordancing are not expected to be interchangeable.

The purpose of this concordance study is to provide score users with assurance that results at specific levels on LANGUAGECERT Academic are comparable with results on IELTS Academic. The results also inform test takers who may be considering different tests and are wondering about which test to take.

Relevant concordance studies conducted between IELTS Academic and other tests within the last fifteen years result in generally strong correlations (Hatch & Lazaraton, 1991), falling in the range of $r = 0.70 - r = 0.78$). Table 1 summarises recent English language test concordance studies. It could therefore be expected that overall correlations between the LANGUAGECERT Academic and IELTS Academic tests would also be in this range, given that both are English language tests intend to cover similar target language use domains.

Table 1. *English language tests concordance studies with IELTS Academic*

Test pair	Authors	Sample size	Overall correlation r
IELTS Academic and PTE-Academic	Clesham & Hughes, 2020	573	0.74
IELTS Academic and PTE-Academic	Elliot et al., 2021	523	0.70
TOEFL iBT and IELTS Academic	ETS, 2010	1,153	0.73
Duolingo English Test and IELTS Academic	LaFlair & Settles, 2019	991	0.78

The current study found a very high overall correlation between LANGUAGECERT Academic and IELTS Academic ($r = .87$). This strong correlation is important as it suggests a substantial similarity in the constructs measured by the two tests. It implies consistency in how the two tests rank test takers according to their language abilities and that both assess similar aspects of academic English proficiency. Scores on one test can be reasonably indicative of performance on the other.

It should be noted that concordance findings are always dependant on the performance of the specific sample of the population assessed and although a large sample size and a robust methodology can allow generalisation, results should be considered as tentative (Knoch, 2021) and test score users are advised to consult concordance tables in combination with additional validation documents and score interpretation frameworks.

LANGUAGECERT Academic – IELTS Academic Concordance study

This concordance study spanned two years, starting in 2022, when the LANGUAGECERT

Academic test was first piloted. It concluded in 2024, having involved more than 1000 test takers, each sitting both the LANGUAGECERT Academic and the IELTS Academic test.

The study and all analyses were designed and supported by the Centre for Research in English Language Learning and Assessment (CRELLA), led by Professor Anthony Green and Dr Johnathan Jones, and the LANGUAGECERT Research team, led by Dr Mike Milanovic and Leda Lampropoulou.

The study compares LANGUAGECERT Academic with IELTS Academic, with a focus on individuals preparing for academic or immigration purposes in English-speaking countries, with a particular attention to Australia. It was designed and conducted to follow good practice in concordance studies (Knoch & Fan, 2024). Findings at each phase were reviewed and approved by the External Validation Panel (formerly known as Concordancing Studies Review panel). The current composition of the panel is as follows:

- Professor Anthony Green, Principal Investigator for the LANGUAGECERT Concordance Study; Professor in Language Assessment, Director of the Centre for Research in English Language Learning and Assessment (CRELLA), University of Bedfordshire, United Kingdom
- Professor Liying Cheng, Professor of Language Education and Assessment, Graduate Faculty Director, Assessment and Evaluation Group, Queen's University, Canada
- Professor Sara Cushing, Professor of Applied Linguistics, ESL Senior Faculty Associate for Assessment of Student Learning, Georgia State University, USA
- Dr Jason Fan, Senior Research Fellow, Deputy Director of the Language Testing Research Centre (LTRC), University of Melbourne, Australia
- Professor James Tognolini, Professor and Director of the Centre for Educational Measurement and Assessment (CEMA), University of Sydney, Australia

During the initial phase of the study, the content comparisons examined the targeted constructs and the language use domains. Having established a high degree of similarity, and the subsequent meaningfulness of linking between test taker performances, the analyses then considered the test scores.

The study was designed so that effects of sequence were counter-balanced, with half the test takers taking the IELTS Academic test first and the other half taking the LANGUAGECERT Academic exam first, within a three-month period, apart from minor exceptions when this was not possible. LANGUAGECERT covered the entire cost of the study without collaborating with the providers of the IELTS test, offering the LANGUAGECERT Academic exam free of charge as well as covering the cost of the IELTS Academic exam for the test taker. Incentivisation was also required to promote motivation, and an additional fee was offered to the test takers depending

on their region.

All participants were required to produce their official IELTS test report form, including marks for total scores and individual skills, while no self-reported scores were accepted. The study followed a single-group design, and no repeat test takers were included. There was a range of test taker nationalities and first languages, involving more than fifteen nationalities and over twenty different first languages, with the majority of test takers having a Chinese or Indian background, broadly representative of the population of interest for the study. The data analysis involved calculating correlations for total scores and scores per skill. Score equation pairs were calculated using equipercentile rankings analysis.

Two complementary approaches were adopted to compare and link the two tests. The first, a qualitative approach, analyses the content of both tests, focusing on their constructs, task types, and the cognitive processes they engage. It also reviews the scoring methods, identifying any limitations that may affect comparability. The second approach, quantitative in nature, focuses on the statistical analysis of test taker performance across the two assessments. Accordingly, this report is divided into two sections: Section 1, "Content Comparisons," which outlines the qualitative analysis, and Section 2, "Statistical Analyses and Results," which presents the quantitative findings.

Pearson correlation coefficients (r) between LANGUAGECERT Academic and IELTS Academic are reported, confirming the strong overall correlation that was expected based on the content comparison, as both tests measure English language proficiency in a similar academic domain. The dataset is then analysed using equipercentile rankings to create score pairs between LANGUAGECERT Academic and IELTS Academic. As a result of the analyses, score pairs between LANGUAGECERT Academic and IELTS Academic emerge. The resulting alignment demonstrates the existence of uneven performance across the four assessed skills, adding to the claim that uneven language competence is the norm (CoE, n.d.; Koizumi et al, 2022; Saville et al., 2021). Even when very strongly linked to the same total score, test taker performance for both tests is seen to vary across skills.

Concluding, LANGUAGECERT asserts that score comparisons across tests, while based on empirical research, are indicative only, and that score users are advised not to rely solely on published score equivalences in making their decisions but they should weigh evidence from additional sources where feasible.

Section 1: Content Comparisons

Summary of LANGUAGECERT Academic and IELTS Academic tests

LANGUAGECERT Academic

The LANGUAGECERT Academic test is an assessment option for test takers seeking to study in English-medium, higher education settings. The test takes approximately 2.5 hours to complete and consists of four parts corresponding to four language skills: Listening (40 minutes), Reading (50 minutes), Writing (50 minutes), and Speaking (14 minutes). The test is designed to assess between the B1 and C2 levels of the Common European Framework of Reference for Languages (CEFR) (Council of Europe 2001). Multiple task types are included to reflect the range of skills used in academic environments, and multiple response types are used for each skill, helping to ensure that no one response type will dominate and potentially bias results (Taylor & Chan, 2015).

The test content analysis measures competences appropriate for academic study in English-medium programmes, such as reading and listening for gist, or detailed understanding of a range of written and audio sources (e.g., academic articles, lectures, podcasts, interviews, and discussions). Additional competences include writing reports, articles and essays for an academic purpose; giving presentations; reading aloud; and taking part in a discussion. Test takers receive a score for each skill (Listening, Reading, Speaking and Writing) along with an overall score and an indication of their CEFR level. In addition, the test taker receives their scores on the LANGUAGECERT Global Scale, a scale which ranges between 0–100 and is aligned to the six levels (A1–C2) of foreign language proficiency described in the CEFR. The alignment between LANGUAGECERT Academic scores on the Global Scale and CEFR levels has also been externally validated (Ecctis, 2023).

The LANGUAGECERT Academic test has been designed for computer-based administration. However, for the concordance study reported here, the administration was initially paper based, as this allowed us to apply strict security measures while operational systems were under development. For the second stage of the study, participants were offered the computer-based version. Regardless of mode of administration, like IELTS, the Speaking component was carried out live with an interlocutor. An equivalence study examined the results obtained on paper-based and computer-based versions of LANGUAGECERT Academic, indicating that the choice of mode did not have a meaningful impact on results.

IELTS Academic

The IELTS Academic test is an established assessment of English language proficiency for individuals who plan to study in an English-speaking setting. The test takes approximately 2

hours and 45 minutes and evaluates the test taker's ability to use and understand English at an academic level. The test is designed to assess test takers between the A1 and C2 levels of the CEFR. It has four modules that assess Listening (30 minutes), Reading (60 minutes), Writing (60 minutes) and Speaking (11–14 minutes). The IELTS Academic and General Training tests are only partly distinct: they have different Reading and Writing components but have the same Listening and Speaking components. In other words, for Speaking and Listening, both the General and Academic tests share the same format and content and the test materials balance more general with more academic English features.

The Listening module requires test takers to listen to recordings of clear, intelligible speech and answer related comprehension questions. The Speaking module involves a face-to-face interview with an IELTS examiner to assess the test taker's English communication skills. The Reading module involves reading three texts and answering comprehension questions. The Writing module consists of two tasks, one involving the description and comparison of data, and the other requiring an essay in response to a prompt. The test taker's overall score on the IELTS Academic test ranges from 0 to 9, with 9 representing the highest level of proficiency. A score is awarded for each skill module as well as for overall performance across all four test modules.

IELTS is available in paper-based and computer formats. For those taking IELTS on computer, the Reading, Writing and Listening modules are completed on a computer, but the Speaking module is conducted face-to-face with an IELTS examiner. For the computer-based test, Speaking is done on the same day as the other components. With the paper version, there can be a 7-day delay between taking the Speaking test and the other modules. Test takers in the concordance study could select their preferred mode of delivery but were not allowed to sit an online proctored exam.

Reporting results and feedback to participants

Results and feedback on performance can have a direct influence on teaching and learning. Similar to the IELTS test, results for LANGUAGECERT Academic are reported both as an overall score and as a score for each of the four language skills. This profile of scores is intended to help language learners to identify areas of strength and areas for improvement. Unlike IELTS, LANGUAGECERT also offers feedback on the productive skill sections (i.e., Speaking and Writing). For Speaking, feedback is given on Task Fulfilment and Communicative Effect, Coherence, Accuracy and Range of Grammar, Accuracy and Range of Vocabulary, and Pronunciation, Intonation and Fluency. Feedback on Writing covers Task Fulfilment, Accuracy and Range of Grammar, Accuracy and Range of Vocabulary, and Organisation.

Providing more detailed information for feedback can help prompt learners to reflect on their performance (Chapelle et al., 2015) and can provide greater opportunity for learners to learn from their current performance so they may move toward their desired level of performance

(Lam, 2021). Feedback can promote self-regulation (Mežek et al., 2022) and can enhance cognitive and emotional engagement (Mayordomo et al., 2022).

Content comparisons

This section reports the content comparisons made between the LANGUAGECERT and IELTS Academic tests. Sections are subdivided by language skill (i.e., Speaking, Listening, Reading, Writing).

Speaking comparison

The following table summarises the Speaking skill comparison between LANGUAGECERT Academic and IELTS Academic. For IELTS, there is a single Speaking module used in both the Academic and General Training versions of the test. In contrast, The LANGUAGECERT Academic Speaking paper has been developed specifically for academic purposes. Key similarities and differences between the two tests are listed below.

Table 2. *Speaking test comparison of LANGUAGECERT Academic and IELTS (Academic and General Training)*

Test	LANGUAGECERT Academic	IELTS (Academic and General Training)
Target Level	B1-C2 / Markscheme covers performance at A1-C2 levels	A1-C2 / Markscheme covers performance at A1-C2 levels
Skills focus	Tasks are designed to elicit speaking skills such as communicating opinions and ideas on a variety of topics (e.g., study skills, news, daily life) and giving personal opinions on contemporary issues. Test takers will also demonstrate their ability to communicate (targeting higher education settings) using a range of functional language to elicit or respond as appropriate, to show the ability to use a wide range of language functions and use of register, to read aloud and answer questions, and to prepare and deliver a presentation in response to a visual stimulus and answer subsequent questions.	Tasks are designed to elicit speaking skills such as communicating personal information, expressing and justifying opinions, explaining, suggesting, speculating, expressing preferences, comparing, summarizing, and narrating.
Skill integration	LANGUAGECERT Academic Speaking entails an integration of speaking, listening, reading, and writing modalities. An examiner (interlocutor) orally explains the tasks and asks questions, requiring the test taker to listen and respond appropriately. In Part 3, the test taker must read a short passage and answer questions from the examiner, and in Part 4	IELTS Academic Speaking entails an integration of speaking, listening, reading, and writing modalities.

Test	LANGUAGECER Academic	IELTS (Academic and General Training)
	test takers must discuss a visual stimulus, such as a chart. In this task, test takers are able to take notes in preparation for the task, engaging writing skills.	
Task description	<p>The exam is delivered in person at a distance by the interlocutor. Speaking tests are recorded.</p> <p>Task 1. (Initial exchange) The examiner introduces himself/herself and confirms the test taker's identity. Test takers give and spell their names and give their country of origin. The examiner then asks up to five general questions on different topics which are expected to be familiar to test takers, such as study skills or daily life. Questions are scaffolded in advance in the examiner's script sheet. After the test taker replies, the examiner responds and/or comments briefly and thanks the test taker before moving to Part 2.</p> <p>Task 2. (Role play) Two situations are presented by the examiner (interlocutor) and test takers are required to respond to and initiate interactions. The examiner begins by explaining that this part of the test is a role play and that the test taker is expected to either start or respond to a situation. The examiner then selects the first role play situation from a prepared list on the examiner sheet, and the interaction continues for approximately two turns before stopping and the next situation is introduced. The examiner selects the second topic from a second list of topic options provided on the examiner sheet. Approximately two turns are given for each interaction, and if there is additional time, an additional situation may be introduced.</p> <p>Task 3. (Read aloud and discussion of passage) The examiner provides the test taker with a Task Sheet which contains a reading passage of approximately 100 words. The examiner allows 30 seconds of preparation time and asks the test taker to read the text out loud. The examiner</p>	<p>The Speaking Test consists of an oral interview between the test taker and an examiner. Speaking tests are recorded.</p> <p>Task 1. (Introduction and interview) The examiner introduces himself/herself and checks the test taker's identity. Then the examiner asks the test taker general questions on some familiar topics such as home, family, work, studies, interests. To ensure consistency, questions are taken from a scripted examiner frame. This part of the test focuses on the test taker's ability to communicate opinions and information on everyday topics and common experiences or situations by answering a range of questions.</p> <p>Task 2. (Long turn) The examiner gives the test taker a task card which asks the test taker to talk about a particular topic, includes points to cover in their talk and instructs the test taker to explain one aspect of the topic. Test takers are given 1 minute to prepare their talk, and are given a pencil and paper to make notes. Using the points on the task card effectively, and making notes during the preparation time, will help the test taker think of appropriate things to say, structure their talk, and keep talking for 2 minutes. The examiner asks the test taker to talk for 1 to 2 minutes, stops the test taker after 2 minutes, and asks one or two questions on the same topic. Part 2 lasts 3-4 minutes, including the preparation time. This part of the test focuses on the test taker's ability to speak at length on a given topic (without further prompts from the examiner), using appropriate language and organising their ideas coherently. It is likely that the test taker will need to draw on their own</p>

Test	LANGUAGECER Academic	IELTS (Academic and General Training)
	<p>further explains that afterwards, the examiner will ask the test taker some questions about the topic. One or more questions, taken from a list on the examiner sheet, is then asked dependent upon time.</p> <p>Task 4. (Presentation) The examiner explains that the test taker will now be asked to give a presentation based on a visual stimulus (e.g. a chart or graph). The test taker has 60 seconds of preparation time, and then must talk about a topic provided by the interlocutor for two minutes. Once the test taker has presented for two minutes or has finished their presentation, the examiner asks follow-up questions for the remaining time.</p>	<p>experience to complete the long turn.</p> <p>Task 3. (Discussion) The examiner and the test taker discuss issues related to the topic in Part 2 in a more general and abstract way and – where appropriate – in greater depth. This part of the test focuses on the test taker’s ability to express and justify opinions and to analyse, discuss and speculate about issues.</p>
Timing	14 minutes approximately (Part 1: 3 minutes; Part 2: 2 minutes; Part 3: 4 minutes; Part 4: 5 minutes)	11-14 minutes (Part 1: 4-5 mins; Part 2: 3-4 mins; Part 3: 4-5 minutes)
Scoring and weighting	<p>Examiners award a raw score of up to 48 points: Task Fulfilment and Communicative Effect (8 points x 2), Coherence (8 points), Accuracy and Range of Grammar (8 points), Accuracy and Range of Vocabulary (8 points), and Pronunciation, Intonation and Fluency (8 points).</p> <p>The exam is delivered in person at a distance by the interlocutor. All tests are recorded. The interlocutor awards the marks for Task Fulfilment and Communicative Effect in real time. A second examiner listens to the recording and awards marks for the other criteria. Discrepancies are checked by a chief examiner. The criterion for Task Fulfilment and Communicative Effect is double weighted.</p> <p>Task Fulfilment and Communicative Effect A measure of the ability to manage the tasks adequately for the level and to communicate successfully with flexibility and naturalness.</p>	<p>Examiners award a band score for each of four criterion areas: Fluency and Coherence, Lexical Resource, Grammatical Range and Accuracy and Pronunciation. The four criteria are equally weighted. Scores are reported in whole and half bands. Detailed performance descriptors have been developed which describe spoken performance at the nine IELTS bands.</p> <p>Fluency and Coherence The ability to talk with normal levels of continuity, rate and effort and to link ideas and language together to form coherent, connected speech. The key indicators of fluency are speech rate and speech continuity. The key indicators of coherence are logical sequencing of sentences, clear marking of stages in a discussion, narration or argument, and the use of cohesive devices (e.g. connectors, pronouns and conjunctions) within and between sentences.</p>

Test	LANGUAGECER Academic	IELTS (Academic and General Training)
	<p>Coherence A measure of the ability to provide coherent responses, particularly over extended speech, and the linking of ideas and contributions.</p> <p>Accuracy and Range of Vocabulary A measure of the ability to vary and demonstrate control of lexis and register as appropriate to the task.</p> <p>Accuracy and Range of Grammar A measure of the ability to vary and demonstrate control of grammatical structures as appropriate to the task.</p> <p>Pronunciation, Intonation and Fluency A measure of the ability to produce the sounds of English in order to be understood with appropriate stress and intonation and maintain the flow of speech.</p> <p>Rating For Speaking, LANGUAGECER uses two raters. The interlocutor awards marks on 'Task Fulfilment and Communicative Effect' as they are communicating directly with the test taker. The other criteria are marked remotely by a second examiner. Collecting marks from two different examiners allows for information to be gathered on performance from two perspectives. The interlocutor/assessor is directly involved in the interaction with the test taker and is therefore in a strong position to judge task completion and communicative effect (also helping prevent prepared answers); however, the interlocutor must concurrently focus on test administration. Having a second examiner who can listen to a recording asynchronously permits greater time and focus to evaluate the more 'analytical' criteria of coherence (i.e., grammar, vocabulary, and pronunciation, intonation and fluency). Enabling examiners to focus on their individual assessment tasks helps promote fair assessment.</p>	<p>Lexical Resource This criterion refers to the range of vocabulary the test taker can use and the precision with which meanings and attitudes can be expressed. The key indicators are the variety of words used, the adequacy and appropriacy of the words used and the ability to circumlocute with or without noticeable hesitation.</p> <p>Grammatical Range and Accuracy This criterion refers to the range and the accurate and appropriate use of the test taker's grammatical resource. The key indicators of grammatical range are the length and complexity of the spoken sentences, the appropriate use of subordinate clauses, and the range of sentence structures, especially to move elements around for information focus. The key indicators of grammatical accuracy are the number of grammatical errors in a given amount of speech and the communicative effect of error.</p> <p>Pronunciation This criterion refers to the ability to produce comprehensible speech to fulfil the Speaking test requirements. The key indicators will be the amount of strain caused to the listener, the amount of the speech which is unintelligible and the noticeability of L1 influence.</p> <p>Rating For Speaking, IELTS uses a single rater. The Speaking is recorded, however, and if there is need for remarking, a second, separate rater will be used for rating.</p>

Test	LANGUAGECERT Academic	IELTS (Academic and General Training)
	Incongruent scores—where a test taker's performance on the task fulfilment criterion is markedly different from performance on the other criteria—are flagged to be reviewed by a third (and deciding) examiner.	
Cognitive processing: Levels of speaking	Conceptualisation	Conceptualisation
	Grammatical encoding	Grammatical encoding
	Phono-morphological encoding	Phono-morphological encoding
	Phonetic encoding	Phonetic encoding
	Self-monitoring	Self-monitoring
	Reciprocal, i.e. direct (face-to-face)	Reciprocal, i.e. direct (face-to-face)
	Planning time allowed	Planning time allowed
Discourse mode	Descriptive, biographical, expository, argumentative	Descriptive, biographical, expository, argumentative
Nature of information	Mix of concrete and abstract	Mix of concrete and abstract
Presentation	Both verbal and non-verbal (e.g. graphs)	Verbal and textual (e.g. a cue card)

Summary of Speaking - Key similarities and distinguishing features

Skill integration

The Speaking components of both tests require an integration of skill modalities, incorporating listening, reading, and writing with speaking. Integrating speaking with reading and listening provides a functional assessment, simulating real-world language use where language skills and processes are simultaneously engaged (Butler et al., 2000), thereby providing greater generalisability to the target language use domain. For both LANGUAGECERT Academic and IELTS Academic, listening is required as the test taker must listen to the interlocutor, interpret aural instructions and questions, and vocalize responses accordingly. In LANGUAGECERT Academic, test takers read and discuss a passage for Task 3, and for Task 4, test takers review a visual stimulus, such as a graph, and present it to the interlocutor. Reading is used in the IELTS exam where participants read instructions for the long monologue task. Note taking is permitted for both tests, meaning the Speaking sections for each can engage all language skills.

Tasks

LANGUAGECERT Task 1 and IELTS Task 1

The initial tasks are similar for LANGUAGECERT and IELTS, discussing concrete, familiar topics; however, less time (by 1-2 minutes) is spent in introductory exchanges in LANGUAGECERT compared to IELTS. Because test takers have more time in the initial IELTS Speaking task, more time is spent on concrete, common, and personally familiar topics. Topic familiarity and concreteness are associated with ease of understanding, theoretically making this part of the speaking section more accessible to lower-level language users. Conversely, in LANGUAGECERT, more focus is given to sections with less personal familiarity, potentially making it more challenging.

LANGUAGECERT Task 2

The two minutes less in LANGUAGECERT's initial task is compensated for in the role play. Role play is a unique element to the LANGUAGECERT Academic intended to engage pragmatic interaction and interactional competence (Lampropoulou, 2022).

LANGUAGECERT Task 3 and IELTS Task 2

The third task in LANGUAGECERT and the second task in IELTS require a combination of reading and speaking. In IELTS, test takers are presented with a card which contains brief notes on a given topic. Test takers are given 1 minute to prepare and are encouraged to take notes, further engaging a multimodal skillset with writing. This contrasts with LANGUAGECERT where test takers are given a passage to read aloud and subsequently discuss.

LANGUAGECERT Task 4 and IELTS Task 3

The final task for both tests is a longer response of approximately 5 minutes. For IELTS, participants build from responses in the previous section, and, with the help of the examiner, are expected to articulate more abstract notions. The final task for LANGUAGECERT Academic is an independent response intended to reflect the participant's ability to present academic information. To aid preparation and organisation, LANGUAGECERT encourages note-making at this stage. Participants are given 60 seconds to view the question prompt (e.g., a graph) and plan their response. After giving a presentation of up to 2 minutes, participants are asked a follow-up question or questions as time permits.

Rating and scoring

Task fulfilment (and rubric misalignment)

LANGUAGECERT Academic and IELTS Academic both explicitly measure coherence, vocabulary,

grammatical range, and pronunciation. There is a difference, however, in how the two tests approach task fulfilment. Task fulfilment assesses a test taker's ability to meet the specific requirements of a task, such as providing a relevant answer to an interlocutor's question. The ability to complete tasks successfully is an important aspect of language proficiency and indicates the test taker's ability to use language for real-world purposes. LANGUAGECERT explicitly includes task fulfilment as an assessment criterion. IELTS Academic may implicitly include task fulfilment as part of the test; however, task fulfilment is not explicitly included in the publicly reported grading criteria. Consequently, there appears to be no direct mechanism for markers to address or report in this regard.

Weighting

In LANGUAGECERT, test takers are awarded a mark from 0-8 for each of five criteria (Task Fulfilment and Communicative Effect, Coherence, Accuracy and Range of Vocabulary, Accuracy and Range of Grammar, and Pronunciation, Intonation, and Fluency). The criterion of Task Fulfilment and Communicative Effect is double-weighted, therefore the maximum raw marks a test taker can be awarded is 48. This necessarily has an influence on analytic scoring. IELTS is evenly weighted across descriptors.

Holistic vs analytic scoring

Both Speaking tests are examined by human raters. IELTS is marked analytically, but scores are reported on a single holistic scale with scores ranging from 0 to 9, in half or full bands. Holistic reporting does not permit detailed feedback on specific areas for improvement and no feedback is provided to test takers other than this single score. LANGUAGECERT Speaking is assessed and reported analytically, with sub-scores for task fulfilment (16 points) along with coherence, vocabulary, grammatical range, and pronunciation (8 points each). LANGUAGECERT provides feedback to the test taker on each of these analytical components. In addition to indicating a more refined explication of performance which may highlight areas of strength and areas for improvement, an analytic approach to scoring promotes transparency.

Number of raters

LANGUAGECERT Speaking includes two raters for Speaking assessment opposed to a single rater used in IELTS Speaking. Analytic scoring increases systematicity and transparency of Speaking ratings; however, rating remains subjective. Raters may have varying standards and can vary in their own consistency. While a second rater can be used to enhance the reliability and validity of the marking (Bejar, 1985), LANGUAGECERT does not implement two raters for each criterion. Instead, LANGUAGECERT splits rating duties between the interlocutor and a second rater. The interlocutor rates "Task Fulfilment and Communicative Effect" while the second rater rates the remaining criteria remotely. Splitting the rating in this way permits the interlocutor to rate in real-time (i.e., is the test taker directly responding to the questions being asked) without the

cognitive burden of balancing test administration with analytic elements of scoring. The remaining criteria being scored remotely enables raters more time for analytic scoring. With IELTS, raters must balance test administration with real-time rating, which can add load and make the examination more challenging for raters (Isaacs et al., 2015). For IELTS, the Speaking section is recorded, and a second rater could be used in cases where remarking is necessary.

In conclusion, the Speaking sections of both LANGUAGECERT Academic and IELTS Academic tests demand an integration of multiple language skills in order to operationalise the assessment of speaking skills. There are, however, some differentiating factors. IELTS Academic has a shared format for both its Academic and General Training variants, while LANGUAGECERT Academic exclusively targets the general academic English language domain. Another difference lies in the marking arrangements. The division of rating responsibilities in LANGUAGECERT Academic enables the interlocutor to focus more on the interaction, potentially leading to a more natural and less pressured conversation for the test taker. In contrast, the IELTS Speaking section requires the examiner to juggle both the administration of the test and the evaluation across all criteria simultaneously. LANGUAGECERT Academic's explicit inclusion of task fulfilment as a criterion, with double weighting, clearly communicates the importance placed on achieving communication. Given these observations, a high positive correlation can be expected between the speaking scores of the two tests, with some variability perhaps at the top scores.

Listening comparison

The following table summarises the Listening skill comparison between LANGUAGECERT Academic and IELTS (there is a single Listening module used in both the IELTS Academic and General Training tests). Key similarities and differences are summarised below.

Table 3. *Listening test comparison of LANGUAGECERT Academic and IELTS (Academic and General Training)*

Test	LANGUAGECERT Academic	IELTS (Academic and General Training)
Target level	B1-C2 / Results are reported on a scale which covers CEFR levels A1-C2, with most items targeting B2 and C1 levels	A1-C2 / Results are reported on a scale which covers CEFR levels A1-C2
Skills focus	<p>Test assesses test taker's ability to:</p> <p>Section 1: identify meaning, purpose and function and understand speaker relationship/context. Conversation completion further tests global comprehension and pragmatic knowledge.</p> <p>Section 2: understand meaning, intention, viewpoint, argumentation and speaker relationship</p>	<p>Test assesses test taker's ability to:</p> <p>Understand details- e.g., listen for names, numbers, and locations and complete a form</p> <p>Section 1: understand concrete, factual information (fill-in form; label a map or diagram)</p>

Test	LANGUAGECER Academic	IELTS (Academic and General Training)
	<p>Section 3: extract key information from a monologue, synthesise and summarise ideas</p> <p>Section 4: follow a discussion between three speakers</p>	<p>Section 2: understand concrete, factual information (complete text; label diagram)</p> <p>Section 3: understand more abstract information (e.g., opinions, arguments, attitudes, inference). Complete comprehension questions or fill in the blank.</p> <p>Section 4: understand more abstract information (e.g., opinions, arguments, attitudes, inference) Complete the summary/ fill in the blank.</p>
Skill integration	<p>LANGUAGECER Academic Listening contains an integration of skills which include reading an answer sheet to identify information required (and preparation time is given to preview), listening to an audio recording, and writing responses.</p> <p>Read forms and texts, predict and identify missing information to listen for, listen and process, then write word(s), label diagrams, or select correct option from MCQ. For gap fill, test takers must select exact words heard, which is a challenge as test takers may provide synonyms based on short term memory and lexical access.</p>	<p>The IELTS Listening test contains an integration of skills which include reading an answer sheet to identify information required (and preparation time is given to preview), listening to an audio recording, and writing responses.</p> <p>Read forms and texts, predict and identify missing information to listen for, listen and process, then write word(s), label diagrams, or select correct option from MCQ. For gap fill, test takers must select exact words heard, which is a challenge as test takers may provide synonyms based on short term memory and lexical access.</p>
Number of items	30	40
Structure and description	<p>4 sections: MCQ (3 tasks) and cloze listening (1 task)</p> <p>Task 1. short conversation completion or continuation. Seven conversations with one 3-option MCQ each - select the option which completes the conversation.</p> <p>Task 2. 5 conversations, 2 MCQ each. Ten total 3-option MCQ across 5 conversations.</p>	<p>4 sections: mix of MCQ, cloze listening, and labelling</p> <p>Task 1. Dialogue (concrete, cloze listening), a conversation between two people set in an everyday social context</p> <p>Task 2. Monologue, set in an everyday social context, e.g., a speech about local facilities</p>

Test	LANGUAGECER Academic	IELTS (Academic and General Training)
	<p>Task 3. cloze listening task featuring a monologue (e.g., a lecture, podcast, narrative, presentation, etc.). There are seven information gaps to fill with up to three words.</p> <p>Task 4. academic discussion (e.g., podcast) or lecture between up to three people. Six 3-option MCQ.</p>	<p>Task 3. Dialogue (academic and training), a conversation between up to four people set in an educational or training context, e.g., a university tutor and a student discussing an assignment</p> <p>Task 4. Lecture, a monologue on an academic subject, e.g., a university lecture.</p>
Timing	<p>40 minutes (with double play) of listening.</p> <p>Question preview time is given for test takers to prepare for the listening. This facilitates contextual understanding as well as focusing attention towards required information.</p>	<p>40 minutes: 30 minutes listening, (paper-based gets 10 minutes to transfer answers from question paper to answer sheet; computer-based gets 2 minutes for answer review). Question preview time is given for test takers to prepare for the listening. This facilitates contextual understanding as well as focusing attention towards required information.</p>
Weighting	<p>Each item is worth one point. Correct answers receive one point while incorrect answer receive zero points.</p> <p>Tasks have different weights (i.e., each input recording is accompanied by a different number of items). Tasks 1 and 3 are worth 7 points each, Task 2 has five dialogues worth 2 points each, totalling 10 points, and Task 4 is worth 6 points. Having fewer points on the more difficult task (Task 4) helps mitigate the effect of its difficulty on the overall skill score.</p>	<p>Each item is worth one point. Correct answers receive one point while incorrect answer receive zero points.</p> <p>Tasks are equally weighted at 10 points each (i.e., each input recording is accompanied by 10 items).</p>
Item density	<p>Given 20 minutes of original recording (omitting double play) and 30 items, one item can be expected approximately every 40 seconds. This is more dense than IELTS. However, the item density is mitigated by double playing the audio, which permits more processing time to confirm responses.</p>	<p>Given 30 minutes of audio recordings and 40 items, one item can be expected approximately every 45 seconds.</p>
Presentation	<p>Listening passages are played twice.</p>	<p>Listening passages are played once.</p>
Cognitive processing:	<p>Factual information</p>	<p>Factual information</p>

Test	LANGUAGECERT Academic	IELTS (Academic and General Training)
Targets	Interpretive information related to context In short tasks, overall understanding of the passage (e.g., Task 1, dialogue completion or continuation)	Interpretive information related to context
Cognitive processing: Levels of listening targeted by items	Word recognition	Word recognition
	Lexical access	Lexical access
	Syntactic parsing	Syntactic parsing
	Identifying the speaker's point	Identifying the speaker's point
	Inference	Inference
	Making referential links	Making referential links
	Inferring the speaker's attitude	Inferring the speaker's attitude
	Integrating meaning to understand key points or meaning in a conversation.	Integrating meaning to understand key points or meaning in a conversation.
Domain	A range of audio sources including academic articles, lectures, podcasts, interviews, discussions.	Recording 1: a conversation between two people set in an everyday social context . Recording 2: a monologue set in an everyday social context , e.g., a speech about local facilities. Recording 3: a conversation between up to four people set in an educational or training context , e.g., a university tutor and a student discussing an assignment. Recording 4: a monologue on an academic subject , e.g., a university lecture
Interaction	Monologic and dialogic	Monologic and dialogic
Discourse mode	Expository, analytical, discursive	Historical/biographical, expository, argumentative
Nature of information	Concrete and abstract, in accordance with an introductory lecture topic	Concrete and abstract, in accordance with an introductory lecture topic
Text length	20 minutes (played twice)	30 minutes

Summary of Listening - Key similarities and distinguishing features

Target levels and timing

Target levels are the first distinguishing feature of the tests. The IELTS Listening test targets CEFR levels A1-C2, which is a broader range than LANGUAGECERT Academic, which targets B1-C2. The IELTS Listening test is correspondingly longer than LANGUAGECERT in terms of number of items (40 items for IELTS compared to 30 items for LANGUAGECERT) and duration (30 minutes of listening for IELTS compared to 20 minutes, single played, for LANGUAGECERT Academic). The IELTS Listening test is shared between the Academic and General Training variants.

Weighting

LANGUAGECERT Academic and IELTS have 30 and 40 items, respectively, with each item weighted equally (one correct response is worth one point). Items are presented in the same order that the relevant information occurs in the audio recordings, offering a degree of scaffolding for the test taker and lessening the cognitive load of the listening task (i.e., less short-term memory is needed). Regarding the weighting of tasks, there is a difference between the tests. Whereas IELTS has equal weighting across its four tasks (10 points each), LANGUAGECERT tasks have different weights, with slightly more weight given to Task 2 (five dialogues worth 2 points each) compared to the others, and slightly less on Task 4 (6 MCQ based on a discussion). As Task 4 is the most challenging, given the nature of the content, this ensures lower-level test takers are not unduly penalised for a task which is beyond their present level.

Item density

LANGUAGECERT has a slightly higher item density, but permits double play, meaning test takers hear information twice. Addressing the lack of double play, IELTS repeats or spells out important information (e.g., a long name or number). Double play also allows time for adjustment as a variety of accents appear across the Listening test in LANGUAGECERT Academic.

Task structure and description

For task structure and description, both LANGUAGECERT and IELTS make use of monologues and dialogues, though there are more dialogues in LANGUAGECERT (12 across the first two tasks opposed to two in IELTS). Dialogues are more extended in IELTS, which tends to make processing more challenging. LANGUAGECERT and IELTS both employ MCQ and cloze listening activities, though in different proportions. In LANGUAGECERT Listening, there are three MCQ tasks and one cloze listening/gap-fill activity. One of the MCQ tasks is unique (there is just one question for each recording) in that the options continue a short conversation, testing global comprehension and pragmatic knowledge. Contrastively, IELTS balances MCQ and cloze listening/gap fill across

tasks. Similarly to LANGUAGECERT, one of the MCQ tasks is unique; however, for IELTS, the unique task employs information matching (e.g., who says what information, or match descriptions to a list of nouns) opposed to conversation continuation. Theoretically, LANGUAGECERT Academic should have an easier structure to follow and less cognitive processing load compared to IELTS.

Nature of information

The LANGUAGECERT Listening test is all set in the academic domain whereas IELTS Listening is shared between the Academic and General Training variants and is, thus, more general English in nature. Descriptively, the nature of information appears more challenging for LANGUAGECERT Academic, with greater emphasis on abstract information compared to IELTS, which is generally more concrete, particularly in Tasks 1 and 2. This may reflect the use of a single Listening test covering both the Academic and General Training tests, as more concrete information may be needed for the General Training than for the Academic test taking population. Initial statistical analysis suggests that the Listening component is easier for IELTS than for LANGUAGECERT, indicating that either the simpler structure has little practical effect on results, or that the nature of the content (and perhaps information density, despite double play) counterbalances the effects of structure on performance. LANGUAGECERT Academic Listening is less directly aligned with IELTS, as it contains less general and familiar topic areas compared with the dual-purpose nature of the Listening test in IELTS Academic. This can be found in the use of lectures as opposed to informational talks in Task 3, or in the use of an academic rather than a topical discussion in Task 4. LANGUAGECERT Listening is, therefore, more closely tied to the academic domain.

Cognitive processing

Higher levels of comprehension, which involve interpreting meaning in context and constructing an understanding of a conversation or lecture, are crucial in academic and professional settings. Both tests assess these higher levels to some degree. LANGUAGECERT and IELTS Listening tasks provide textual support with visual aids, such as a torn piece of paper with writing on it or a notepad with a pencil; however, IELTS additionally offers diagrammatic aids to fill in (e.g., map labelling). This offers additional context for learners, and labelling tasks may benefit or hinder test takers depending on their aptitude for spatial awareness. Some individuals may be disproportionately affected by a given item or task type and having multiple task types helps to mitigate this risk. However, MCQ is a very commonly employed and widely known item type, and LANGUAGECERT's predominant use of MCQ may help to facilitate understanding of the task, permitting participants to focus more on the content than on how to respond to the task.

Comparison limitations: listening difficulty is impacted by numerous factors which have not been described here as we do not have access to the IELTS test specifications. For instance, speech rate, accent, and lexical and grammatical complexity in the audio recordings can all impact a

listener's ability to accurately comprehend speech. However, assuming certain industry standards are followed (e.g., speech rate of 150 words/minute; targeting a percentage of B2-level vocabulary for academic texts), LANGUAGECERT and IELTS may be expected to be similar in these respects.

In conclusion, the comparison between LANGUAGECERT Academic and IELTS Academic Listening tests reveals several significant similarities, which suggest that the listening scores will be positively correlated. Both tests assess listening skills in English, including the ability to understand main ideas and specific factual information, recognise the opinions, attitudes, and purpose of speakers, and follow the development of an argument. Both tests are aligned with the Common European Framework of Reference for Languages (CEFR), even though they target slightly different ranges within the framework. Both tests include academic content and aim to assess the test taker's ability to understand spoken English in academic settings. At the same time, there are certain factors that can be expected to affect the strength of the correlation. IELTS Academic offers a broader range of CEFR levels and a shared format suitable for both Academic and General Training candidates. LANGUAGECERT Academic focuses more on the academic context with a slightly higher item density and the feature of double play. Given these considerations, one might expect a moderate to high positive Pearson correlation coefficient between the scores on the LANGUAGECERT Academic and IELTS Listening tests, indicating that candidates who score well on one test are likely to score well on the other.

Writing comparison

The following table summarises the Writing skill comparison between LANGUAGECERT Academic and IELTS Academic. Key similarities and differences are summarised below.

Table 4. *Writing test comparison of LANGUAGECERT Academic and IELTS Academic*

Test	LANGUAGECERT Academic	IELTS Academic
Target Level	B1-C2 / Markscheme covers performance at A1-C2 levels	A1-C2 / Markscheme covers performance at A1-C2 levels
Response format and Genre	Task 1: Continuous writing; Information transfer from visual or textual input (e.g., a chart, graph, or table) Task 2: A longer piece of continuous writing; Essay	Task 1: Continuous writing; Information transfer from multiple non-verbal inputs Task 2: A longer piece of continuous writing; Essay
Task description	The LANGUAGECERT Academic Writing test is designed to assess a wide range of writing skills, including how well test takers: <ul style="list-style-type: none"> • write a response appropriately • organise ideas • use a range of vocabulary and grammar accurately 	The IELTS Academic Writing test is designed to assess a wide range of writing skills, including how well test takers: <ul style="list-style-type: none"> • write a response appropriately • organise ideas • use a range of vocabulary and grammar accurately

Test	LANGUAGECER Academic	IELTS Academic
	<p>Task 1. Test takers explain a visual input (e.g., a graph) in 150-200 words with the intended reader specified. Test takers might describe and explain data, express a stance, opinion, justification, or argument in accordance with the provided prompt.</p> <p>Task 2. Test takers explain opposing views of a given topic and express their own opinion in 250 words. Test takers read two opposing opinions, discuss both views, and write their personal perspective.</p>	<p>In Task 1 test takers are presented with a graph, table, chart or diagram. Test takers are asked to describe, summarise or explain the information in their own words. This might involve describing and explaining data, describing the stages of a process or how something works, or describing an object or event.</p> <p>In Task 2 test takers are asked to write an essay in response to a point of view, argument or problem.</p>
Domain	Academic, social	Academic, social
Purpose	<p>Task 1:</p> <ul style="list-style-type: none"> To demonstrate the ability to understand and synthesise visual or textual inputs To show the ability to write a report, argument or article using a written, graphic or visual input with the intended reader specified expressing stance, opinion, justification, argumentation. <p>Task 2:</p> <ul style="list-style-type: none"> To write a formal piece of writing for a specified reader which may compare and contrast, persuade, argue, hypothesise, evaluate, analyse, or present solutions. 	<p>Task 1:</p> <ul style="list-style-type: none"> To transfer information from multiple inputs To collate different pieces of information in order to describe, summarise or explain the information. <p>Task 2:</p> <ul style="list-style-type: none"> To write a persuasive essay To defend or attack a particular argument or opinion, compare or contrast aspects of an argument, and give reasons for the argument.
Timing	50 minutes. No explicit instruction is provided to divide time, but an expected word count is indicated for Task 1 (150-200 words) and Task 2 (250 words).	60 minutes. Test takers should spend 20 minutes on Task 1, and 40 minutes on Task 2. Test takers need to manage their own time.
Text length of expected response	Task 1: 150-200 words Task 2: 250 words	Task 1: at least 150 words Task 2: at least 250 words
Weighting	Task 1: 40% Task 2: 60%	Task 1: 33.3% Task 2: 66.6%
Skills assessed	In both tasks, test takers are assessed on their ability to write a response which is appropriate in terms of content, the organisation of ideas, and the accuracy and range of vocabulary and grammar. Task fulfilment is an explicit part of the	In both tasks, test takers are assessed on their ability to write a response which is appropriate in terms of content, the organisation of ideas, and the accuracy and range of vocabulary and grammar. Task fulfilment is an

Test	LANGUAGECERT Academic	IELTS Academic
	marking criteria.	explicit part of the marking criteria.
Cognitive processing	<p>Task 1:</p> <ul style="list-style-type: none"> • Macro-planning: goal setting and task representation • Transforming ideas from verbal and non-verbal inputs • Organising ideas • Translating • Micro-planning • Monitoring and revision • Comprehending non-graphic task instructions • Comprehending (and interpreting) the components of graphs • Re-presenting or re-producing the non-graphic and graphic information as continuous discourse in written form in English as a foreign language <p>Task 2:</p> <ul style="list-style-type: none"> • Macro-planning: goal setting and task representation • Generating ideas • Organising ideas • Translating • Micro-planning • Monitoring and revision 	<p>Task 1:</p> <ul style="list-style-type: none"> • Macro-planning: goal setting and task representation • Transforming ideas from non-verbal inputs • Organising ideas • Translating • Micro-planning • Monitoring and revision • Comprehending non-graphic task instructions • Comprehending (and interpreting) the components of graphs • Re-presenting or re-producing the non-graphic and graphic information as continuous discourse in written form in English as a foreign language <p>Task 2:</p> <ul style="list-style-type: none"> • Macro-planning: goal setting and task representation • Generating ideas • Organising ideas • Translating • Micro-planning • Monitoring and revision
Discourse mode (rhetorical task)	Descriptive, expository, argumentative/persuasive	Descriptive, expository, argumentative/persuasive
Scoring approach	Analytic	Analytic (reported holistically)
Scoring and marking	<p>Test takers are assessed on their performance on each task by trained examiners according to four criteria: Task Achievement, Accuracy and Range of Grammar, Accuracy and Range of Vocabulary, and Organisation (Coherence).</p> <p>Scores are reported both holistically and analytically. A holistic score is provided using the Global Scale (out of 100), and additional feedback is provided based on performance on individual marking criteria (i.e., task achievement, accuracy</p>	<p>Test takers are assessed on their performance on each task by certificated IELTS examiners according to the four criteria of the IELTS Writing Test Band Descriptors (task achievement/response, coherence and cohesion, lexical resource, grammatical range and accuracy).</p> <p>Scores are reported holistically in whole and half bands.</p> <p>Between two and four examiners mark IELTS Writing assessments.</p>

Test	LANGUAGECER Academic	IELTS Academic
	<p>and range of grammar, accuracy and range of vocabulary, and organisation (coherence).</p> <p>Marking Marking is double blind between two examiners. If a large discrepancy exists between the two examiners, the writing response is passed to a third senior examiner whose marks are final.</p>	<p>Marking Marking is done between two and four examiners for accuracy and fairness.</p>

Summary of Writing - Key similarities and distinguishing features

Timing

Less time is devoted to the Writing section for LANGUAGECER Academic (50 minutes) compared to IELTS Academic (60 minutes). The expected word counts are similar across tasks, however, and test takers taking the LANGUAGECER exam will likely experience more time pressure compared to IELTS test takers.

Tasks

Task 1

The initial tasks are similar for LANGUAGECER Academic and IELTS Academic, where test takers are prompted to discuss a visual input (e.g., graph, table, diagram) for a minimum of 150 words. The writing is formal in register. LANGUAGECER Task 1 differs from IELTS Task 1 in how it supports responses. LANGUAGECER provides a response scaffold through the prompt. For instance, a chart is provided with the instruction to write a report which describes the main trends, give reasons for the trends, and predict likely changes over a given period of time. Contrastively, IELTS Academic instructions are more open to interpretation on the part of the test taker: for example, asking test takers to summarise a chart and “make comparisons where relevant”.

Task 2

The second task for both tests is a longer response, argumentative essay, with a recommended minimum of 250 words. As with Task 1, the writing is formal in register. For LANGUAGECER, test takers are asked to read two opposing opinions, summarise them, and offer their own perspective on the topic. For IELTS, one opinion is offered, and the test takers must explain the extent to which they agree or disagree with the given perspective. The opposing view (e.g., “children who are brought up in families that do not have large amounts of money are better prepared to deal with the problems of adult life than children brought up by wealthy parents” is

inferred (IELTS online test preparation). Because only one perspective is offered to an argumentative topic, the prompts may more strongly elicit a personal or emotional reaction than LANGUAGECERT, which may influence results.

Rating and scoring

Task fulfilment (and rubric alignment)

The Writing section rubrics for LANGUAGECERT and IELTS are more strongly aligned than in the Speaking section, with a direct correspondence in each rubric component. In the Writing section markers score coherence, grammatical and vocabulary accuracy and range, but distinct for LANGUAGECERT Academic compared to IELTS Academic, task fulfilment is an explicit criterion.

Weighting

For both tests, greater scoring weight is given to Task 2, the extended writing task, than Task 1. For LANGUAGECERT, Task 1 is given a weight of 40% of the total marks for Writing, and Task 2 is given a weight of 60% of total marks. For IELTS, slightly more weight is given to Task 2, with approximately 2/3rds of the weight allocated to Task 2 compared to 1/3rd for Task 1. Given the weighting, test takers in IELTS are instructed to divide their time accordingly across tasks, with 20 minutes devoted to Task 1 and 40 minutes to Task 2. On the test form, LANGUAGECERT does not explicitly advise test takers on which component should receive more time, though it does state the suggested number of words that should be provided for each task.

Holistic vs analytic scoring and reporting

Both Writing tests are examined by human markers. IELTS Writing is assessed analytically and reported holistically on a scale of 0–9, whereas LANGUAGECERT Writing is assessed and reported analytically, with sub-scores (worth a maximum of 8 points each) for Task Achievement, Accuracy and Range of Grammar, Accuracy and Range of Vocabulary, and Organisation (Coherence). LANGUAGECERT provides feedback on these analytical components, whereas IELTS simply reports a band score. In addition to indicating a more refined explication of performance which may highlight areas of strength and areas for improvement, an analytic approach to scoring promotes transparency in scoring.

Number of markers

Both LANGUAGECERT and IELTS attempt to ensure reliability and validity of marking through multiple markers. LANGUAGECERT and IELTS employ a minimum of two markers for the Writing tasks. LANGUAGECERT will use a third (who holds the final decision) if discrepancies exist, while IELTS can use up to four.

In conclusion, the Writing sections of LANGUAGECERT Academic and IELTS Academic share key similarities in their aims to evaluate candidates' ability to articulate ideas in written English, emphasizing formal register, critical thinking, and analytical engagement with prompts. Both tests underscore the importance of the extended writing task through higher weighting, reflecting the value placed on argumentation and coherence in academic contexts. Despite a slight difference in timing, the fundamental objectives and evaluation criteria are aligned, as rating scales for both tests are heavily derived from the relevant CEFR scales. Given these similarities, a moderate to high positive correlation could be expected between scores in the Writing sections of LANGUAGECERT Academic and IELTS Academic.

Reading comparison

The following table summarises the Reading skill comparison between LANGUAGECERT Academic and IELTS Academic. Key similarities and differences are summarised subsequently.

Table 5. *Reading test comparison of LANGUAGECERT Academic and IELTS Academic: Task Features*

Test	LANGUAGECERT Academic	IELTS Academic
Target Level	B1-C2 / Results are reported on a scale which covers CEFR levels A1-C2, with most items targeting B2 and C1 levels	A1-C2 / Results are reported on a scale which covers CEFR levels A1-C2
Skills focus	<p>To test students' ability to comprehend academic texts and to extract important information from those texts. The test is designed to assess a wide range of reading skills, including how well test takers:</p> <ul style="list-style-type: none"> • read for the general sense of a passage • read for the main ideas • read for detail • understand vocabulary used in academic texts, identify synonyms and use vocabulary in context • understand lexico-grammatical features in academic texts • understand inferences and implied meaning • understand how meaning is built up in discourse and show awareness of text organisation and discourse features • understand long complex texts, including discourse, opinion, purpose, argumentation, exemplification, comparison and contrast, cause and effect and locate specific information 	<p>To test students' ability to comprehend academic texts and to extract important information from those texts. The test is designed to assess a wide range of reading skills, including how well test takers:</p> <ul style="list-style-type: none"> • read for the general sense of a passage • read for the main ideas • read for detail • understand inferences and implied meaning • recognise a writer's opinions, attitudes and purpose • follow the development of an argument
Task	There are 30 questions. A variety of	There are 40 questions across three

Test	LANGUAGECER Academic	IELTS Academic
description	<p>question types is used, chosen from the following types: multiple choice, gap fill, identifying information, identifying writer's views/claims, matching information to source texts, sentence and text completion, synonym identification.</p> <p>Task 1a. (4-option synonyms) Six sentences written in an academic style with one word highlighted. Test takers choose a synonym for highlighted word from a list of four words.</p> <p>Task 1b. (3-option cloze) An authentic academic text that may include academic ideas, arguments and opinions with five words removed. Test takers choose the correct word from a choice of three to full each gap</p> <p>Task 2. (Gapped sentences) An academic text with 6 sentences removed. Test takers choose from 8 sentences to complete the text.</p> <p>Task 3. (Short text comprehension with multiple matching) Test takers match seven questions to four texts. Texts may be reviews, reports, articles, journals, opinion pieces, etc. with a linked theme, but with a different purpose.</p> <p>Task 4. (Long text comprehension with 4-option MCQ) An extended text (e.g., narrative, descriptive, explanatory, expository, biographical, instructive) with 6 MCQs.</p>	<p>parts (three passages). A variety of question types is used, chosen from the following types: multiple choice, identifying information, identifying writer's views/claims, matching information, matching headings, matching features, matching sentence endings, sentence completion, summary completion, note completion, table completion, flow-chart completion, diagram label completion, short-answer questions. Sometimes one-word answers are required, sometimes a short phrase, and sometimes simply a letter, number or symbol.</p> <p>Mainly receptive, some limited writing involved in short answer questions, but only brief answers are required; no more than a given number of words. Test takers lose marks for incorrect spelling and grammar.</p>
Number of items	30	40
Timing	50 minutes to answer 30 questions on 7 passages (across 4 tasks.)	60 minutes to answer a total of 40 questions on 3 passages. Individual tasks are not timed.
Weighting	All items equally weighted. Each correct answer receives one mark. Tasks have different weights (i.e., each text is accompanied by a different number of items). Scores out of 30 are converted to the Global Scale out of 100 for reporting purposes.	All items equally weighted. Each correct answer receives one mark. Scores out of 40 are converted to the IELTS Academic 9-band scale. Scores are reported in whole and half bands

Test	LANGUAGECER Academic	IELTS Academic
Cognitive processing Goal setting	<p>Expeditious reading: local (scan/search for specifics)</p> <p>Expeditious reading: global (skim for gist/search for key ideas/detail)</p> <p>Careful reading: local (understanding sentence)</p> <p>Careful reading: global (comprehend main idea(s)/overall text(s))</p>	<p>Expeditious reading: local (scan/search for specifics)</p> <p>Expeditious reading: global (skim for gist/search for key ideas/detail)</p> <p>Careful reading: local (understanding sentence)</p> <p>Careful reading: global (comprehend main idea(s)/overall text(s))</p>
Cognitive processing Levels of reading	Word recognition	Word recognition
	Lexical access	Lexical access
	Syntactic parsing	Syntactic parsing
	Establishing propositional meaning (cl./sent. level)	Establishing propositional meaning (cl./sent. level)
	Inferencing	Inferencing
	Building a mental model	Building a mental model
	Creating a text level representation (disc. structure)	Creating a text level representation (disc. structure)
	Creating an intertextual representation (multi-text)	Creating an intertextual representation (multi-text)

Table 6. Reading test comparison of LANGUAGECERT Academic and IELTS Academic: Features of the Input Text

Test	LANGUAGECERT Academic	IELTS Academic
Word count	<p>There are 7 passages to read, plus questions, across 4 sections. Given provided texts, test takers must read approximately 2750-2800 words, including source text and questions.</p> <p>Long and short texts are provided to elicit different skills. Longer texts permit an assessment of skimming and scanning for information, and discriminate between important and secondary details. Shorter texts are used to assess vocabulary (synonyms) and lexico-grammatical knowledge.</p>	<p>Three different passages to read, each with accompanying questions. Officially test takers have to read 2,150 - 2,750 words in total. There are three sections to the IELTS Academic Reading test, and each contains one long text.</p> <p>Green, Ünalı and Weir and (2010) analysed 42 texts making up 14 IELTS reading tests. The passages in their study contained 854 words on average (maximum 1063 words, minimum 589 words).</p>
Average sentence length	16.32 words per sentence on average across all sentences.	21.89 words per sentence on average across all sentences.
Domain	Academic	Academic
Discourse mode	Narrative, descriptive, explanatory, expository, biographical, instructive	Historical/biographical, expository, argumentative
Nature of information	Mix of concrete and abstract	Mostly concrete
Presentation	Verbal (textual)	Both Verbal (textual) and Non-verbal (i.e., graphs)
Lexical Level; further criteria	<p>The cumulative coverage reaches 95.6% at the K3 level. 75.2 K1, 12.3 K2, 8.1 K3.</p> <p>Lexical Density 0.58</p>	<p>The cumulative coverage reaches 92% at the K3 level. 76.4 K1, 11.36 K2, 3.26 K3.</p> <p>Lexical Density 0.57</p>
Readability	Flesch-Kincaid Grade Level 10.27	Flesch-Kincaid Grade Level 12.64
Topic	A range of topics which may reasonably reflect introductory coursework and texts at the university level are employed. Topics represented in sample material include controversial technology in sport, animal domestication, language preservation, and wildlife and deep-sea mining.	A broad range of subject areas were represented among the 42 IELTS texts examined by Green et al. (2010) with the categories of <i>Social studies</i> (10 or 11 texts), <i>Engineering & technology</i> (6 or 7) and <i>Business & administrative studies</i> (4 or 5) emerging as the most popular topic areas for the test.
Text genre	The seven reading texts are appropriate to the general academic genre, including vocabulary and topic content. Longer texts mimic introductory texts (e.g., textbook excerpts, academic reports) which cover a range of concrete and abstract information.	Three reading texts with a variety of question types. The kinds of text used in IELTS introduce academic topics to a general audience, often in the form of articles sourced from newspapers or magazines presenting research findings to a general audience. These include self-contained reports on developments in science and

Test	LANGUAGECERT Academic	IELTS Academic
	Given the nature of the test and the lack of assumed technical knowledge, text excerpts are potentially less difficult than course content which is both extended and domain specific.	technology and overviews of academic debates. While IELTS passages are at a level of difficulty appropriate to university study, they are not as challenging as some of the texts encountered in the more linguistically demanding areas such as the law textbook analysed by Green et al. (2010).

Summary of Reading - Key similarities and distinguishing features

Target levels and timing

Targeted proficiency levels are the first distinguishing feature of the tests as they impact test design. The IELTS Academic Reading test targets CEFR levels A1-C2, a broader range than LANGUAGECERT Academic, which targets B1-C2. The IELTS Academic Reading test is correspondingly longer than LANGUAGECERT Academic in terms of number of items (40 items for IELTS Academic compared to 30 items for LANGUAGECERT Academic) and duration (60 minutes of Reading for IELTS Academic compared to 50 minutes for LANGUAGECERT Academic).

Weighting

Reading components for LANGUAGECERT Academic and IELTS Academic have 30 and 40 items, respectively, with each item weighted equally (one correct response is worth one point). Items are presented in the same order that the relevant information occurs in the passages, offering a degree of scaffolding for the test taker and lessening the cognitive load of the reading task (i.e., less short-term memory is needed). Regarding the weighting of tasks, there is a difference between the tests. Whereas IELTS Academic has equal weighting across its four tasks (10 points each), LANGUAGECERT tasks have different weights across tasks. Tasks 1A and 1B combine for 11 points, Task 2 is worth 6 points, Task 3 is worth 7 points, and Task 4 is worth 6 points. As Task 4 is the most challenging, given its length and the nature of the content, this ensures lower-level test takers are not unduly penalised for a task which is beyond their present level.

Task description

LANGUAGECERT has more reading passages compared to IELTS, but fewer questions. Longer texts are associated with greater levels of difficulty; however, an increased number of passages may also increase difficulty. LANGUAGECERT includes components which more directly assesses elements of vocabulary and grammar (Task 1). The structure of LANGUAGECERT is more predictable than IELTS, as tasks and item types are largely uniform in the Reading test. Contrastively, IELTS Reading can have different numbers of items per passage and there are at least 14 question types which may be given (as shown in the Comparison Table).

LANGUAGECERt's test, therefore, may permit greater scaffolding and targeted preparation. However, it does so at the cost of response type coverage. Further, offering different item types can help prevent the performance of individual test takers being unduly influenced by specific item types, whether due to relative strength or weakness.

Input texts

Input texts are similar in genre as both tests target the academic domain. However, the texts in IELTS Academic were longer and somewhat more difficult overall compared with LANGUAGECERt Academic, with a readability score of 12.64 compared to 10.27 (readability is associated with grade level according to the US school system), and fewer words at the K3 level or below (92% for IELTS Academic compared to 96% for LANGUAGECERt Academic).

Cognitive processing

Though IELTS Academic offers more diversity in option types, both of the tests appear to engage similar levels of processing, including word recognition, lexical access, syntactic parsing, establishing propositional meaning, inferencing, building a mental model, creating a text level representation of discourse structure, and creating an intertextual representation across texts.

In conclusion, the Reading sections of IELTS Academic and LANGUAGECERt Academic are designed to evaluate candidates' proficiency in handling academic texts, albeit with distinct approaches in terms of target proficiency levels, test duration, and item distribution. In comparing the two tests, several aspects suggest that the LANGUAGECERt test could present certain challenges. While IELTS Academic targets a wider range of CEFR levels (A1-C2), LANGUAGECERt Academic focuses on B1-C2 levels, concentrating its content on higher proficiency levels. Despite having fewer items, the weighting of tasks in LANGUAGECERt Academic is distinct, with varying points allocated to different sections. Moreover, the LANGUAGECERt Academic test incorporates more reading passages than IELTS Academic. The inclusion of more passages could elevate the difficulty by requiring test takers to adjust to different texts more frequently, thereby taxing their cognitive resources more heavily.

Conclusion to Content Comparison Section

Compared in detail in the sections above, the constructs and their assessment across the two examinations – LANGUAGECERT Academic and IELTS Academic – show high degrees of similarity. Mainly, they are both four-skill tests with the intention of measuring language competence in order to predict the ability to operate successfully in an academic setting. Having said that, there are several differentiating factors.

The LANGUAGECERT Academic test is based in the academic context across all four skill components whereas IELTS Academic focuses entirely on the academic domain in the Reading and Writing tests, but is more General English in nature in the Listening and Speaking tests.

The objectively marked components of the IELTS Academic tests have more items in order to cover the full range of CEFR levels (A1-C2) whereas the LANGUAGECERT Academic tests for Reading and Listening focus exclusively on the B1-C2 levels, primarily targeting B2 and C1 as these are the levels typically required for entrance onto English-language medium courses of study.

There is a wider range of task types for the receptive tests of IELTS Academic while the LANGUAGECERT Academic examination has a fully consistent test format, with each sub-skill test having the same task formats and number of items in each test version. The Reading and Listening sub-skills tested show a very high degree of similarity.

The marking criteria for the tests of Writing and Speaking show a high degree of similarity. One exception is the 'Task Fulfilment and Communicative Effect' criterion for LANGUAGECERT Academic Speaking. This is not explicitly measured as a separate criterion in IELTS. In terms of the scoring of tasks, the tests assess a very similar range of levels. LANGUAGECERT Academic has a two-examiner model for the Speaking test meaning marks are given on a test taker's performance from two different perspectives, whereas IELTS Academic Speaking has, as a standard, a one-examiner model. The marking of both Writing tests features at least two examiners who award marks independently.

The outcome of the content comparison suggests that the two tests address the same domains and share numerous similarities. It appears likely that results on the two tests will correlate positively. Notable differences do emerge at the skill level, as highlighted in the conclusions for each skill component. Among others, these differences include such features as content, question density, and specificity of focus. Throughout the study, careful consideration was given to the potential impact of these differences on the interpretation of findings.

Section 2: Statistical analyses and results

The goal of the concordance study was to identify how performance on the LANGUAGECERT Academic (LCA) exam relates to performance on IELTS Academic. IELTS Academic was selected as a comparison for the present study as it is a similar measure of English language proficiency targeting general academic English as a language use domain, it has been mapped to several different frameworks, is generally accepted for admissions and immigration purposes, and has publicly available results, published annually, which could readily be used for comparison.

The analyses herein provide a descriptive overview of the test and its subskills, report how test takers perform on the test and compare that performance to performance on IELTS Academic. Correlations are provided to demonstrate the quantitative relationship between performances on the two tests. For the purposes of this report, the key output is a concordance table which sets out how bands on the IELTS Academic correspond to scores on LANGUAGECERT Academic, and conversely, how scores on LANGUAGECERT Academic translate to band scores on IELTS Academic. This is followed by a brief discussion of population invariance, looking at the extent to which subpopulation performances are aligned with the results of the concordance tables.

The results and conclusions of the concordance study are reliant on the sample they are based on, and it is therefore instructive to first identify the sample as it compares to the larger population of potential test takers. In this section, the test taker characteristics and their performance on the concordance study are described and, where possible, compared with the characteristics and performance of the larger IELTS Academic test taking population and the wider population of individuals needing to prove their English competency for higher education admission and immigration purposes to English-speaking countries, and to Australia in particular.

Test taker demographics

Sample population

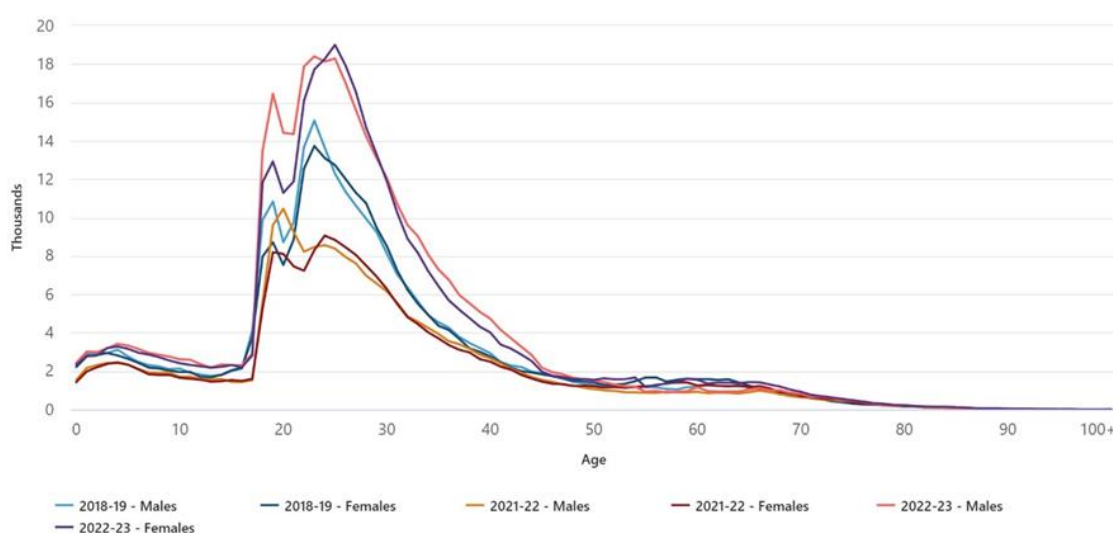
The sample population was carefully selected to closely mirror the demographic of interest for the study, i.e., individuals with profiles similar to IELTS Academic test takers, who wish to use their test score to pursue migration for academic or educational purposes in higher learning institutions where English is the medium of instruction, such as universities in Australia. Nationalities and first language cover a wide range; expected English language competency spans mainly B1 to C1 CEFR levels. Gender split is relatively equal. Eligible test takers needed to also adhere to strict guidelines as to which test they needed to take first, the time period within which both tests had to be taken, as well as providing LANGUAGECERT with access to their official IELTS Test Report Form. This resulted in a concordance dataset comprising the performances of 1008 unique test takers taking both the LANGUAGECERT Academic and IELTS Academic tests and

on which all subsequent analyses are based.

Age

Regarding age, the average test taker age was 24.64 with a standard deviation of 6.31 years, meaning most test takers were between 18 and 31 years old. The most recent publicly available IELTS Academic statistics for 2022 (<https://ielts.org/researchers/our-research/test-statistics>) do not include data for test taker age. For that reason, the sample's appropriacy was evaluated against data retrieved from the Australian Bureau of Statistics (Overseas Migration, 2022-23 financial year (<https://www.abs.gov.au/statistics/people/population/overseas-migration/latest-release> | Australian Bureau of Statistics for annual updates). In 2022-23, the median age of migrant arrivals was 27 and the modal age was 25. This age range is also typically expected for undergraduate and postgraduate study (House, 2010). Consequently, the concordance study sampling population is representative of Australia's current migration data in terms of age.

Figure 1. *Overseas migrant arrivals - age and sex*



Source: Australian Bureau of Statistics, Overseas Migration 2022-23 financial year

As the publicly available IELTS Academic statistics do not show performance by age, comparison of results by age group across LANGUAGECERT Academic and IELTS Academic could not be achieved.

Gender

With regard to gender, the concordance study sample strongly corresponded with the test population of interest. There were 602 (59.72%) female participants, 404 (40.08%) male, and 2 (0.2%) participants who did not disclose their gender. This is a slightly higher proportion of females compared to the IELTS Academic 2022 statistics, which report 52% female and 48% male. Mean IELTS Academic scores in this study closely match the means for IELTS Academic

total scores in the published data. Gender split and performance for both tests are shown in the tables below, demonstrating a close match to the total scores means of the published data.

Table 7. IELTS Academic mean total scores by gender for the study sample compared to published IELTS Academic data (2022).

Gender	Test takers	% of sample	IELTS 2022 data - % of sample	LCA total score mean	IELTS total score mean	IELTS 2022 data - total score mean
Female	602	59.72	51.78	64.50	6.33	6.28
Male	404	40.08	48.22	62.29	6.16	6.22
Undisclosed	2	0.20	N/A	66.50	6.75	N/A
Total	1008	100	100	63.62	6.26	N/A
Female $n = 602$, male $n = 404$, undisclosed $n = 2$, total $n = 1008$.						

Table 8. IELTS Academic mean skill component scores by gender for the study sample compared to published IELTS data (2022).

Gender	Concordance study	IELTS 2022
	Listening	
Female	6.40	6.51
Male	6.40	6.52
	Reading	
Female	6.50	6.26
Male	6.20	6.20
	Writing	
Female	5.90	5.95
Male	5.90	5.86
	Speaking	
Female	6.00	6.17
Male	6.00	6.06

Nationality

Regarding nationality, the concordance study aimed to be as representative of the test population as possible. The largest non-native English speaking populations, who may require an English language test for migration purposes, arrive from India, China, and the Philippines (Australian Bureau of Statistics, 2024). Half (51%) of international student arrivals originated from China and India, with additional significant numbers coming from Nepal, Colombia, the Philippines, Thailand and Pakistan. Concordance test taker nationalities are presented in table 9.

Table 9. *Test taker Nationality*

Nationality	Test takers	% of sample
Chinese	471	46.73
Indian	260	25.79
Iraqi	89	8.83
Spanish	38	3.77
Greek	30	2.98
Thai	22	2.18
Turkish	18	1.79
Colombian	11	1.09
Pakistani	10	0.99
Iranian	7	0.69
Japanese	5	0.50
Brazilian	3	0.30
Italian	3	0.30
Saudi Arabian	3	0.30
Syrian	3	0.30
Ukrainian	3	0.30
French	2	0.20
South Korean	2	0.20
Other	28	2.78
Total	1008	100.00

Tables 10 and 11 below further compare mean scores for overall scores and subskills split by nationality, demonstrating how the study population aligns with the published IELTS Academic test population statistics.

Table 10. *IELTS Academic mean total scores by nationality for the study sample compared to published IELTS data (2022).*

Nationality	Concordance study	IELTS Academic 2022
	Overall	
All (n=1008)	6.3	6.3
China (n=471)	6.2	6.1
India (n=260)	6.3	6.2
Colombia (n=11)	6.7	6.6
Thailand (n=27)	6.0	6.1
Pakistan (n=10)	5.8	6.2

Table 11. IELTS Academic mean skill component scores by nationality for the study sample compared to the published data (2022).

Nationality	Concordance study	IELTS 2022	Concordance study	IELTS 2022
	Listening		Reading	
All (n=1008)	6.4	6.5	6.4	6.2
China (n=471)	6.4	6.1	6.7	6.4
India (n=260)	6.6	6.6	6.1	6.0
Colombia (n=11)	6.9	6.7	6.7	6.9
Thailand (n=27)	6.1	6.3	6.0	6.1
Pakistan (n=10)	6.1	6.5	5.4	6.1
	Writing		Speaking	
All (n=1008)	5.9	5.9	6.0	6.1
China (n=471)	5.9	5.8	5.7	5.6
India (n=260)	6.1	5.9	6.1	6.1
Colombia (n=11)	5.9	6.0	6.7	6.6
Thailand (n=27)	5.6	5.7	5.7	5.8
Pakistan (n=10)	5.7	5.9	5.5	6.2

The study population closely resembles the test population of interest in terms of performance. For both overall scores and subskills, test taker performance is within 0.1-0.7-point difference. The maximum difference is less than an IELTS Academic band (0.7) and this is only in two subskills and a very small sample size (i.e., Pakistan, n=10).

First language

Table 12. *Test taker first language background*

First language (mother tongue)	Test takers	% of sample
Chinese	476	47.22
Punjabi	96	9.52
Kurdish	95	9.42
Malayalam	75	7.44
Spanish	51	5.06
Tamil	41	4.07
Greek	30	2.98
Hindi	22	2.18
Thai	22	2.18
Telugu	20	1.98
Turkish	19	1.88
Farsi	8	0.79
Urdu	8	0.79
Arabic	5	0.50
Japanese	5	0.50
French	4	0.40
Portuguese	4	0.40
Italian	3	0.30
Ukrainian	3	0.30
English	2	0.20
German	2	0.20
Korean	2	0.20
Other	15	1.49
Total	1008	100.00

Self-reported proficiency

During the registration stage, participants were asked to indicate their language proficiency indicating a CEFR level. As it could not be expected that all participants would be familiar with the CEFR levels, and/or able to self-estimate accurately, they were asked to take an online placement test which LANGUAGECERT has available on its website and which consists of 25 multiple-choice questions that assess grammar and vocabulary (<https://www.LANGUAGECERT.org/en/preparation/quiz>). They would then select one of the six CEFR levels, which also included a description of the Can-do statements as per the CEFR Table 1 (CEFR 3.3) Global scale: Common Reference levels (CoE, 2001).

Self-reported proficiency levels had a distribution across all CEFR levels. Among the test takers who self-estimated their English proficiency, there were 49 A1, 21 A2, 84 B1, 206 B2, 204 C1, and 66 C2. The bulk of proficiency levels being at B1-C2 reflects the target range of the

LANGUAGECERT Academic test.

Sequence and Exam interval

All participants took both tests (LANGUAGECERT Academic and IELTS Academic), and a counterbalanced design was used to mitigate potential order effects. Counterbalancing occurred such that test takers were split into two administration groups: Group 1 took LANGUAGECERT Academic first and IELTS Academic second; Group 2 took IELTS Academic first and LANGUAGECERT Academic second, within a three-month period, apart from minor exceptions that were then examined and found to be statistically non-significant. A perfect balance was sought but was not always possible due to test taker schedules and availability. The final sequence of administration was 459 taking IELTS Academic first and 549 taking LANGUAGECERT Academic first.

An independent samples t-test was conducted to examine the effect of exam sequence on test taker performance. For LANGUAGECERT Academic test taker performance, overall mean performances were marginally lower for those taking LANGUAGECERT Academic first ($M = 63.17$, $SD = 12.85$) compared to those who took the IELTS Academic first ($M = 64.12$, $SD = 11.97$); however, this was not found to be statistically significant, $t(994.311) = 1.216$, $p = .224$. Similar results were found for IELTS Academic test taker performance. Overall mean performances were marginally lower for those taking IELTS Academic first ($M = 6.25$, $SD = .85$) compared to those who took the LANGUAGECERT Academic first ($M = 6.26$, $SD = .91$). This again was not found to be statistically significant, $t(1008) = .265$, $p = .791$. For this dataset, sequence, or taking one exam before the other, did not play a significant role in test taker performance.

Restricting the time between administrations helps minimise the possibility of test takers' language proficiency improving from one test to another, potentially biasing results. This study was designed so that the interval between test takers attempting both tests was as small as possible, and that it should not be greater than three months (Dorans et al., 1997).

The targeted three-month window was largely maintained, with the average time between the two sittings being 38 days, and the majority ($n = 868$, or 86%) of the test takers taking both tests within 90 days (inclusive). Despite best efforts to adhere to this timeframe, however, the study period of 2022-2024 coincided with the enforcement of COVID-19 pandemic-related regulations in certain countries, notably affecting the schedule. Specifically, the Chinese government imposed severe restrictions in various cities, including Beijing and Shanghai, from November 2022 to February 2023, disrupting exam schedules due to prohibitions on access to testing facilities for candidates and staff. Additionally, this led to difficulties for participants in securing available IELTS Academic slots following a nearly three-month suspension of exams. As a result, a minority of participants ($n = 99$, or 10%) experienced slight delays, but managed to undertake the second test between 91 and 120 days (inclusive). A further 4% of participants ($n = 41$)

encountered more significant delays, completing the second examination between 121 days and 203 days of the first.

We retained test takers who took the tests more than 90 days apart because:

- statistical analyses showed that the difference in test taker performance on the first test and second test taken was not statistically significant regardless of whether tests were taken within 90 days or beyond 90 days (see Appendix A – Exam Interval analysis),
- the average interval between administrations of 38 days is relatively short compared to other large scale language testing concordance reports (e.g., Clesham and Hughes, 2020),
- the additional 140 test takers provided valuable data and helped achieve the targeted 1,000 participants for data analysis and
- any sensitivity to duration between test administrations would be mitigated through the study's counterbalanced design.

Familiarity

As LANGUAGECERT Academic is a relatively new language test, a difference of familiarity was expected between the LANGUAGECERT Academic test and the more established IELTS Academic. Familiarity (and subsequently motivation) may have an effect on test taker performance, and therefore reasonable attempts to control and document it were made. It was controlled to the extent possible by making test information materials available to all test takers. Test takers were provided with familiarisation packages for both tests covering test format and task types, practice papers and a prerecorded webinar. They were also advised and reminded to access the materials prior to sitting the exam.

Prior to the test, familiarity was measured using a 4-option Likert-type scale (not familiar, a little familiar, familiar, very familiar) (see Appendix B). Of respondents who indicated their familiarity with LANGUAGECERT Academic (n = 625), the mode response was “a little familiar”. Test takers were more familiar with IELTS Academic, as test takers who indicated their familiarity (n = 631) had a mode response of “familiar”.

The effect of test familiarity on test taker performance was examined in terms of mean scores. Test takers are seen to have a mean score of 63-64 out of 100, regardless of their familiarity with the LANGUAGECERT Academic test. For IELTS Academic, being familiar or very familiar with the test indicated a slightly higher mean score of 6.41 against 6.03 of those with a little or no familiarity, but this was not found to be statistically significant. The impact of familiarity on test performance is presented in the tables below.

Table 13. *Test taker familiarity with the LANGUAGECERT Academic test at exam registration*

LC Academic	n	Mean	SD
No response	383	62.82	11.74
Not /A little familiar	405	64.19	12.38
Very / Familiar	220	63.88	13.75

Table 14. *Test taker familiarity with the IELTS Academic test at exam registration*

IELTS Academic	n	Mean	SD
No response	377	6.13	0.86
Not /A little familiar	127	6.03	0.94
Very / Familiar	504	6.41	0.86

Test takers' reasons for taking the test

Using the same questionnaire, test takers were asked about their primary reasons for taking each test (see Appendix B). Reasons were provided by 659 test takers concerning the LANGUAGECERT test while 622 test takers gave reasons for taking the IELTS Academic exam. Participants could choose as many reasons as appropriate to their circumstances, hence the total numbers are not relevant. The results are presented in the table below. Using the certificate for “a higher education extended course (more than three months)” was the main reason for taking both tests (26% for LCA and 38% for IELTS) followed by “other educational purposes” and “personal reasons”.

Table 15. *Test takers' reasons for taking the LANGUAGECERT Academic and the IELTS Academic tests*

I intend to use my test results for...	LANGUAGECERT Academic	IELTS Academic
higher education extended course	26.1%	38.1%
higher education short course	2.0%	1.7%
other education purposes	13.9%	11.7%
registration as a healthcare professional	1.1%	2.2%
professional registration (NOT healthcare)	0.6%	1.1%
employment purposes	7.6%	8.4%
immigration purposes	4.7%	6.6%
personal reasons	16.5%	10.6%
other purpose /concordance	5.4%	0.3%

Score distribution summary statistics

Shown in the overall distribution statistics in Table 16 below, summary statistics indicate a normal distribution for the data.

Table 16. *Distribution statistics for LCA and IELTS Academic*

	Mean	SD	Skew	Kurtosis	Min	Max
LANGUAGECERT Academic	63.62	12.46	-.106	-.077	22	96
IELTS Academic	6.26	0.88	-.059	-.144	3	8.5

Sample size = 1008

Individual means and standard deviations were also calculated for each subskill, shown in Table 17 below.

Table 17. *Subskill descriptive statistics comparison between LANGUAGECERT Academic and IELTS Academic*

	Overall	Listening	Reading	Writing	Speaking
LANGUAGECERT Academic	63.62 (12.46)	60.48 (17.18)	63.07 (17.25)	61.96 (11.07)	68.43 (12.76)
IELTS-A	6.26 (0.88)	6.42 (1.28)	6.38 (1.31)	5.93 (0.64)	6.01 (0.88)

Sample size = 1008. Indices include mean followed by standard deviation in brackets.

Test taker performance analysis

Correlations

Correlations were conducted between LANGUAGECERT Academic and IELTS Academic. Taylor and Chan (2015) reported correlations between several tests of English for Academic Purposes. They found overall correlations between selected tests (i.e. CAE, CPE, TOEFL iBT, OET and PTE-A) and IELTS Academic ranged from 0.73 to 0.87. A strong, positive correlation between tests (i.e. above 0.7) suggests that performing well on one test would translate to performing well on the other, while performing poorly on one test generally corresponds with performing poorly on the other.

Because both tests in this study measure English language proficiency, a strong overall correlation can be expected. We found a correlation between LANGUAGECERT Academic and IELTS Academic of 0.87, matching the highest overall correlation between assessments as reported by Taylor and Chan (2015). This relationship is visually represented in Figure 2.

Figure 2. Overall score relationship between LANGUAGECERT Academic and IELTS Academic

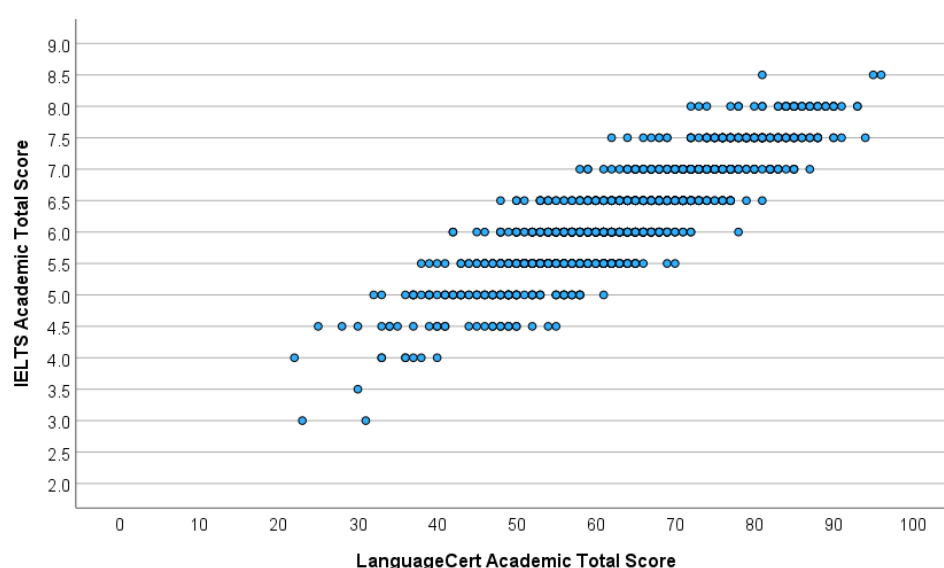


Figure 2 shows the linear relationship between tests, where a high score on LANGUAGECERT Academic is associated with a high score in IELTS Academic, while a low score on LANGUAGECERT Academic is associated with a low score in IELTS Academic.

Correlations between scores for component skills were also examined, with results shown in Table 18 alongside Overall performance.

Table 18. Correlations between LANGUAGECERT Academic and IELTS Academic

Overall <i>r</i>	Reading <i>r</i>	Writing <i>r</i>	Listening <i>r</i>	Speaking <i>r</i>
.87	.76	.71	.71	.71

Note: *r* = correlation. All correlations were statistically significant at the $p < .001$ level. Sample size = 1008.

The strong correlation for Overall performance ($r=.87$) indicates the two tests measure similar underlying abilities (Knoch, 2021). The correlations for individual subscales ($r > .7$) suggest that LANGUAGECERT exams perform in accordance with the benchmark language test designed for similar purposes. This meets or exceeds correlations between alternative tests and IELTS Academic as reported in the studies in table 1 in this report.

Comparing test scores

It was important to establish what a given score on a LANGUAGECERT Academic test corresponds to for performance on an IELTS Academic test, and to use those scores to develop a summary concordance table for ease of reference. Two common linking methods were considered: linear regression and equipercentile rankings. With linear regression, it is not necessary to have data for each point in a given test score range and it can be calculated with smaller samples (Johnson & Oswald, 2017), making it desirable for studies with relatively few participants. Consequently, linear regression was implemented in the earlier stages of the concordance project when the sample was less robust. The use of linear regression for linking,

however, is limited because it is unidirectional. It provides different score equivalences depending on which test is used for prediction. Specifically, a score on IELTS Academic could be predicted from scores on LANGUAGECERT Academic exams, but scores on IELTS Academic could not be used to predict scores on LANGUAGECERT Academic.

As such, interpreting figures in the concordance table would be less intuitive for potential test users. The equipercentile linking method addresses this issue as it offers a bidirectional link between scores; a score on IELTS Academic can be used to interpret a score on LANGUAGECERT Academic and vice versa, making the concordance table easier to understand.

Therefore, upon reaching sufficient levels for equipercentile calculations (i.e., $n \geq 1000$), equipercentile linking was implemented. Equipercentile calculations were made based on a single group, counterbalanced design. Analyses were conducted using the statistical analysis program, R (R Core Team, 2021) and the R package, “equate” (Albano, 2016). A common approach to equipercentile linking employs pre-smoothing to the data, such as loglinear pre-smoothing—an established method of reducing irregularities in the data due to sample size (Hanson et al., 1994; Kolen & Brennan, 1995). This was considered; however, results were similar or identical to when no smoothing was applied, and it was determined that the small changes made were insufficient to justify altering the data with smoothing (Fairbank, 1985). Results have been consolidated into the concordance table below (Table 19).

Table 19. Overall alignment table for LANGUAGECERT Academic and IELTS Academic performance. See recommendations for interpretation and use of linkage results for test users below.

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.

Skill comparison

Following this overall view, concordance in each of the four skills was considered. The same approach was adopted as with overall scores. Linking was implemented to rankings with no smoothing. Analyses were conducted using the statistical analysis program, R (R Core Team, 2021) and the R package, “equate” (Albano, 2016). On close inspection of performance across

the skills, some anomalies in the distributions were observed in the subtests, and particularly in Reading and Listening where more test takers received IELTS bands 8 and 9 than corresponding scores on the LANGUAGECERT Academic Listening and Reading components. We also noticed that this skewed distribution was not reflected in the IELTS Speaking and Writing tests. At the same time, it was observed that the standard deviations of the IELTS Speaking and Writing tests were narrower than their LANGUAGECERT Academic counterparts indicating that there was a greater tendency for IELTS Speaking and Writing markers to regress to the mean. Such differences are discussed at greater length later in the report but it is important to note that these differences were not considered to compromise the equipercentile scaling described in this section.

As expected, the relationships between the skills present a more variable picture. In the subjectively marked skills, the relationship between LANGUAGECERT Academic and IELTS Academic is reasonably close and within expectations. It is only at the lowest end of the proficiency continuum that the writing test results diverge more significantly. This may in part be attributable to the paucity of data at the lowest levels. Given that both LANGUAGECERT Academic and IELTS Academic marking schemes draw heavily on CEFR level descriptions a close relationship is to be expected.

There is greater divergence between IELTS Academic and LANGUAGECERT Academic in the objectively marked skills. LANGUAGECERT Academic appears to be more demanding. This was expected to be the case to some extent with the listening test, given that it is aimed at an academic audience where the IELTS Academic listening test is of a more general nature. We note in this context that IELTS Academic has significantly more high performing test takers scoring band 8 or 9 as opposed to LANGUAGECERT Academic. It is not clear why this should be the case, but it affects the relationship between the two tests at the top end of the scale. Interestingly, these levels of higher performance are not reflected in IELTS Academic Speaking or Writing. The IELTS Academic Reading test has a more academic orientation, similar to LANGUAGECERT Academic. However, we note a similar pattern of high performing test takers in the IELTS Academic sample as we did with the IELTS Academic listening test. This high level of performance is not reflected in the tests of Speaking and Writing.

The concordance data suggests ranking pairs for each skill between LANGUAGECERT Academic and IELTS Academic as indicated in the tables below.

Table 20. *Listening skill alignment table for LANGUAGECERT Academic and IELTS Academic performance. See recommendations for interpretation and use of linkage results for test users below.*

IELTS Academic Listening	LANGUAGECERT Academic Listening	n-size of study sample at this level	Standard Error
4.5	35-40	40	1.48
5.0	41-48	108	1.13
5.5	49-56	147	1.03
6.0	57-61	152	1.03
6.5	62-66	135	0.99
7.0	67-72	109	1.30
7.5	73-79	100	1.28
8.0	80-88	78	1.31
8.5	89-94	77	1.20
9.0	95-100	25	1.51

Table 21. *Reading skill alignment table for LANGUAGECERT Academic and IELTS Academic performance. See recommendations for interpretation and use of linkage results for test users below.*

IELTS Academic Reading	LANGUAGECERT Academic Reading	n-size of study sample at this level	Standard Error
4.5	36-43	59	1.51
5.0	44-53	125	1.04
5.5	54-59	161	0.91
6.0	60-64	126	1.03
6.5	65-70	133	0.98
7.0	71-76	98	1.10
7.5	77-82	93	1.08
8.0	83-88	68	1.35
8.5	89-96	82	1.11
9.0	97-100	34	1.58

Table 22. Writing skill alignment table for LANGUAGECERT Academic and IELTS Academic performance. See recommendations for interpretation and use of linkage results for test users below.

IELTS Academic Writing	LANGUAGECERT Academic Writing	n-size of study sample at this level	Standard Error
4.5	33-44	18	3.34
5.0	45-55	91	1.01
5.5	56-63	221	0.56
6.0	64-70	389	0.36
6.5	71-77	199	0.48
7.0	78-83	65	0.71
7.5	84-88	11	1.06
8.0	89-92	2	9.00
8.5	93+	1	n/a
9.0	n/a	0	n/a

Table 23. Speaking skill alignment table for LANGUAGECERT Academic and IELTS Academic performance. See recommendations for interpretation and use of linkage results for test users below.

IELTS Academic Speaking	LANGUAGECERT Academic Speaking	n-size of study sample at this level	Standard Error
4.5	44-53	34	1.79
5.0	54-61	107	0.92
5.5	62-69	211	0.71
6.0	70-75	275	0.49
6.5	76-81	171	0.60
7.0	82-86	109	0.70
7.5	87-88	45	1.16
8.0	89-92	19	1.52
8.5	93-98	8	1.22
9.0	99-100	3	3.38

Recommendations for interpretation and use of linkage results for test users

Score users, for example institutions who use certain test scores for decisions about test takers, are advised that score comparisons across tests, while based on empirical research, are estimates only and should be treated with caution for the following reasons:

- Tests differ, sometimes significantly, in the ways information about English language ability is elicited and assessed. Score comparisons are only meaningful to the extent that the tests are measuring the same ability or skill.
- Tests often differ in the length of the reporting scales used (for example one test may report on a 6-point scale and another on a 100-point scale). As a result, a one-to-one mapping of scores from one test to another is rarely possible.
- The choice of concordance study methodology may produce variations in results.
- The populations of test takers may differ (e.g., with respect to ages, nationalities, language backgrounds of test takers) from the population used in the research that generated the score equivalences.
- The sample sizes used for comparing scores from different tests are generally small across all levels/ranges, especially at extreme ends of the scale.
- Score concordance results are generally more robust for proficiency levels with larger numbers of test takers.
- Large Standard Errors show that score equivalences are particularly imprecise at certain points on the ability scale.

Because the score comparisons presented in the score comparison table are indicative only, 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.

Population invariance

Population invariance—where similar linking results should be found regardless of the subpopulation that the linking function is applied to—has been posited as a fundamental condition of linking and equating studies (Dorans & Holland, 2000). It was therefore relevant to not only identify the concordance of scores between LANGUAGECERT Academic and IELTS Academic, but to investigate how results with subpopulations corresponded with these findings. Though subdividing the study to investigate subgroup performance necessarily reduces the sample size and can complicate subsequent analyses (Brennan, 2008), preliminary calculations were conducted to establish processes going forward. Two measures were taken to explore invariance, one at the test level, and one at the item level. Invariance at the test level will be outlined here by using the previously described equipercentile method on key subgroups.

An initial population invariance was explored by looking at the largest subgroups within the study, namely, gender (female and male) and nationality (Chinese and non-Chinese). Results are summarised in the tables 24 to 28 below.

Table 24. Population Invariance - Overall

IELTS Academic Score	Female	Male	Chinese	All Other Nationalities
4.5	37-45	40-46	37-47	39-45
5	46-53	47-53	48-54	46-52
5.5	54-60	54-59	55-60	53-59
6	61-67	60-65	61-67	60-66
6.5	68-72	66-71	68-72	67-71
7	73-80	72-79	73-81	72-78
7.5	81-86	80-87	82-89	79-86
8	87+	88+	90+	87+

Table 25. Population Invariance - Speaking

IELTS Academic Score	Female	Male	Chinese	All Other Nationalities
4.5	43-52	46-53	44-54	43-49
5	53-62	54-60	55-62	50-60
5.5	63-70	61-69	63-71	61-67
6	71-75	70-76	72-78	68-74
6.5	76-81	77-81	79-85	75-79
7	82-86	82-86	86-88	80-85
7.5	87-88	87	89-93	86-87
8	89+	88+	94+	88+

Table 26. Population Invariance - Listening

IELTS Academic Score	Female	Male	Chinese	All Other Nationalities
4.5	36-41	33-39	38-43	27-37
5	42-49	40-46	44-52	38-44
5.5	50-57	47-54	53-59	45-53
6	58-62	55-59	60-63	54-59
6.5	63-68	60-64	64-69	60-64
7	69-73	65-69	70-73	65-71
7.5	74-81	70-75	74-80	72-77
8	82+	76+	81+	78+

Table 27. Population Invariance – Reading

IELTS Academic Score	Female	Male	Chinese	All Other Nationalities
4.5	37-44	35-43	34-41	36-44
5	45-53	44-53	42-51	45-54
5.5	54-59	54-59	52-58	55-60
6	60-63	60-66	59-62	61-66
6.5	64-69	67-72	63-68	67-72
7	70-75	73-77	69-74	73-78
7.5	76-81	78-82	75-80	79-84
8	82+	83+	81+	85+

Table 28. Population Invariance – Writing

IELTS Academic Score	Female	Male	Chinese	All Other Nationalities
4.5	35-44	31-44	36-44	32-44
5	45-56	45-55	45-56	45-55
5.5	57-64	56-62	57-63	56-63
6	65-71	63-69	64-71	64-69
6.5	72-77	70-77	72-78	70-75
7	78-84	78-82	79-85	76-81
7.5	85-89	83-87	86+	82-88
8	90+	88+	n/a	89+

The above results indicate a general overlap across groups; however, variance is present and should continue to be investigated as subpopulations become more robust with increased sample sizes. As previously indicated, more data are needed at the extremes (e.g., 8 and above and 4 and below), and partitioning the dataset to investigate subgroup performance has reduced the number of data points further. Though more data is needed, the strong reliability shown in both tests further indicates support for invariance within the sample. As Dorans and Holland (2000, pp 300-301) state, “whenever the reliabilities of X and Y (the two tests, LANGUAGECERT Academic and IELTS Academic, in this case) are both high, “near population invariance” is expected to hold for a wide range of subpopulations.” LANGUAGECERT Academic was found to have strong reliability across language skills (McDonald’s $\omega = .88$), supporting potential invariance.

Concordance results discussion

The qualitative comparisons between LANGUAGECERT Academic and IELTS Academic demonstrate that the tests are very similar in construct and content, with LANGUAGECERT Academic being more focussed on the academic domain in terms of the Listening and Speaking tests. Statistically, there are robust correlations between LANGUAGECERT Academic and IELTS Academic overall (.87) and by skill (all >.70) suggesting that the two tests are measuring language abilities in very similar ways.

Speaking

In terms of the separate language skills that results are reported on, for LANGUAGECERT Academic Speaking – each main score with a full descriptor on the rating scale (8, 6, 4, 2 and 1) is tied to a CEFR level. That is, 8 is designed to reflect C2 level CEFR Can-Do descriptors, 6 is tied to C1, 4 to B2 and so on. An analysis of the marking criteria for IELTS Academic and LANGUAGECERT Academic shows that they are similar in nature and descriptions of performance, and so, assuming reliable interpretation of the mark scheme, similar results would be expected. This is, indeed, the case with a slightly wider spread of scores for LANGUAGECERT Academic. Therefore, from a qualitative and quantitative perspective the two Speaking tests align very well. It is noticeable that the concordance study candidature achieved their highest mean scores on LANGUAGECERT Academic on the Speaking component. This may be because, even though the LANGUAGECERT test is less well-known, the Speaking test is the one component where a test taker is actively engaged in an interaction with a human interlocutor and so performs to the best of their abilities. This is likely to have some effect on the overall concordance despite best efforts to minimise any lack of motivation through effective design and incentivisation. It is also worth noting the paucity of IELTS Speaking performances at levels 8 and above (only 2.8% of the test takers), which contrasts with the percentage of test takers achieving a total score of 8 and above overall (30.2%).

Writing

In terms of the Writing tests, for LANGUAGECERT Academic each score with a full descriptor on the markscheme for LANGUAGECERT Academic (8, 6, 4, 2 and 1) is tied to a CEFR level. That is, 8 reflects CEFR C2 level Can-Do descriptors, 6 is tied to C1, 4 to B2 and so on. An analysis of the marking criteria for IELTS Academic and LANGUAGECERT Academic shows that they are closely aligned (having the same roots in the CEFR Can-do descriptors). Assuming consistent interpretation of the rating scales, similar results would be expected. This is, indeed, the case with a slightly wider spread for LANGUAGECERT Academic scores. Therefore, from both the qualitative and quantitative perspective the two Writing tests appear to align well. A closer examination of the Writing section shows that IELTS Academic scores are slightly compressed compared to LANGUAGECERT Academic scores. Only 3 out of the 1008 test takers in the study achieved an IELTS Academic writing score of 8 or above.

Figure 3. IELTS Academic Writing scores distribution

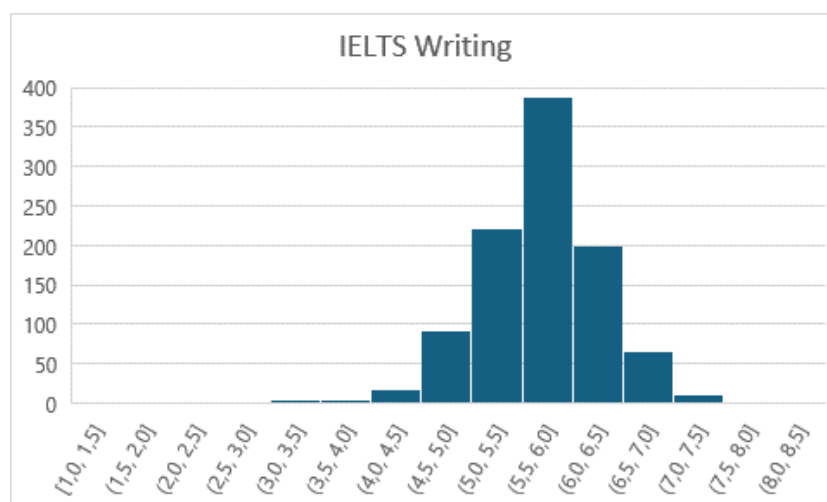
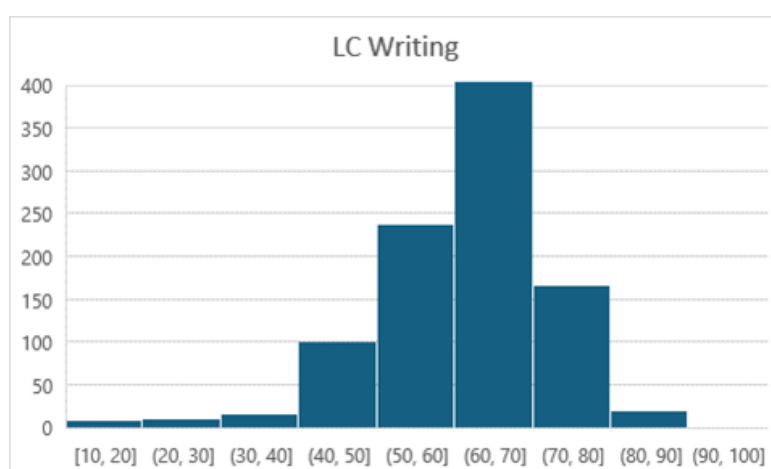


Figure 4. LANGUAGECERT Academic Writing scores distribution



Listening

From the concordance study, the receptive skills tests show more difference in terms of test taker results across the two examinations. There are certain features of the comparison between Listening scores in LANGUAGECERT Academic and IELTS Academic that should be noted. In the data used for this study, the LANGUAGECERT Academic Listening results show a classic normal distribution, whereas the IELTS Academic distribution has a normal distribution with a slight positive skew illustrating a larger proportion of test takers scoring very highly (bands 8.0-9.0) compared to overall IELTS Academic scores.

Figure 5. IELTS Academic Listening scores distribution

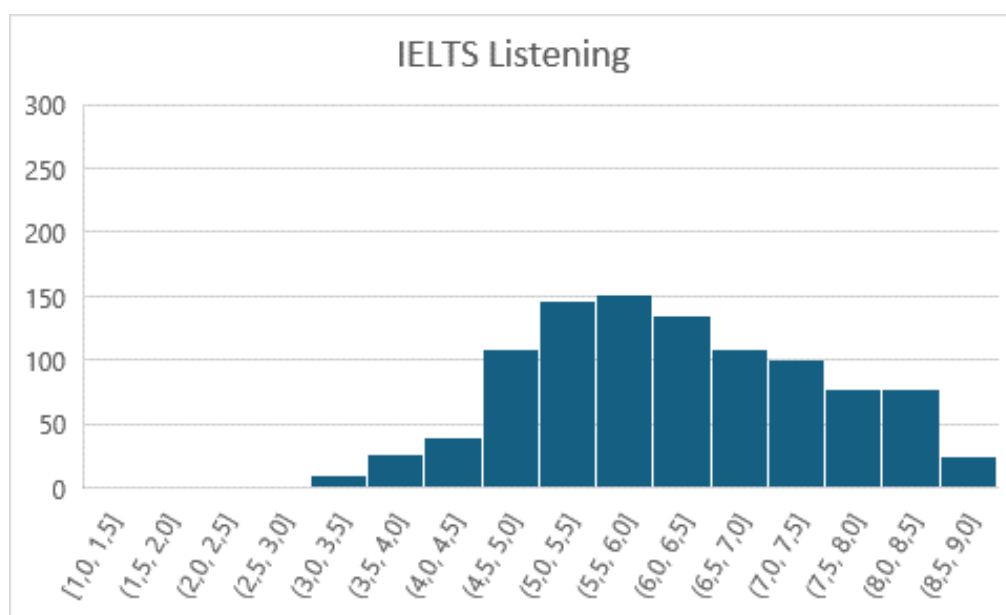
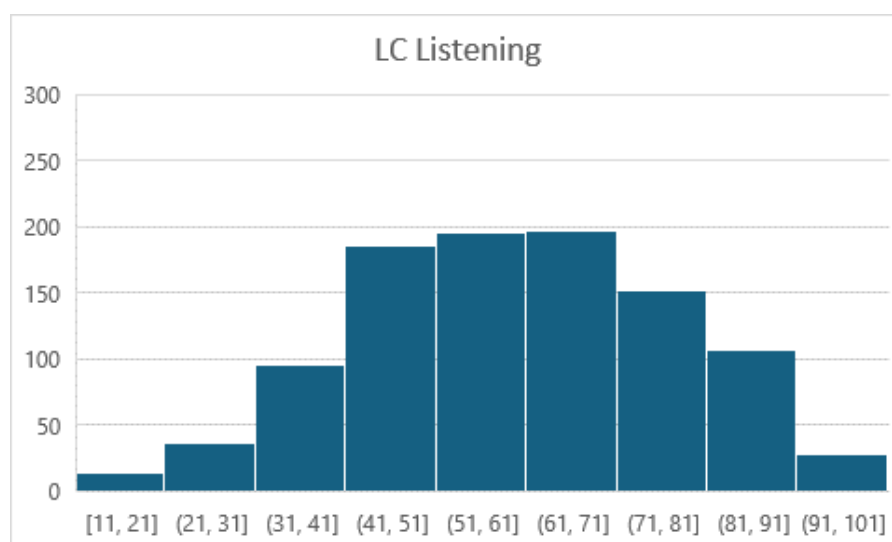


Figure 6. LANGUAGECERT Academic Listening scores distribution



The LANGUAGECERT Academic Listening items used in the concordance study have good item statistics, with test reliability figures above .80, and item facilities almost all in the target range (.30 to .85) and high levels of discrimination. These statistics indicate tests that have performed

well. However, in addition to the differences in score distributions, there was some evidence in the concordance study that the LANGUAGECERT Academic Listening test, perhaps due to its more academic nature, was slightly more difficult for test takers than the IELTS Academic Listening test.

Reading

An analysis of the LANGUAGECERT Academic Reading tests presents a similar picture to Listening, although with less marked differences. The distribution for IELTS Academic Reading again shows a slightly positively skewed normal distribution showing a larger proportion of candidates scoring very highly (bands 8.0 to 9.0) and higher than their overall IELTS Academic scores. The LANGUAGECERT Academic Reading tests administered in the concordance study also produced strong reliability figures (around .85), item facilities were almost all in the target range (.30 to .85) and with high levels of discrimination. These statistics again show tests that performed well. However, there was some evidence in the concordance study that the LANGUAGECERT Academic Reading test was slightly more difficult for test takers than the corresponding IELTS Academic Reading test.

Conclusion to Statistical analyses and results section

Concluding this analysis, it is worth pointing out that alignment in English language testing is not a singular, unidimensional process; rather, it is a continuous, complex and intricate procedure that demands sustained attention and evolves over time. Continuous validation involves regularly assessing the test against established criteria to verify its accuracy and effectiveness, whereas alignment requires consistent adjustments to ensure that test content aligns with current language standards and accurately reflects the skills and competencies being measured. These processes are essential for maintaining the validity and fairness of English language tests over time, acknowledging the evolving nature of language and the diversity of the test taker population. LANGUAGECERT advises test score users to treat all score comparison tables as indicative only, and not to rely solely on published score equivalences in making their decisions. They should weigh evidence from additional sources where feasible. LANGUAGECERT's concordancing studies programme will continue beyond the point of conclusion of this study with the view to a) confirming the healthy performance of our test within the different contexts and b) determining concordance with comparable English language tests, such as TOEFL iBT, PTE Academic and others as appropriate in the context and for the purpose the test is used.

References

- Australian Government, Department of Home Affairs. (2023). *Migration strategy*. Retrieved from <https://immi.homeaffairs.gov.au/programs-subsite/migration-strategy/Documents/migration-strategy.pdf>
- Bejar, I. I. (1985). Test speededness under number-right scoring: An analysis of the Test of English as a Foreign Language. *ETS Research Report Series*, 1985(1), i-57.
- Butler, F. A., Eignor, D., Jones, S., McNamara, T., & Suomi, B. K. (2000). *TOEFL 2000 speaking framework*. Princeton, NJ: Educational Testing Service.
- Chapelle, CA, Cotos, E., & Lee, J. (2015). Validity arguments for diagnostic assessment using automated writing evaluation. *Language testing*, 32(3), 385-405.
- Clesham, R., & Hughes, S. R. (2020). concordance report PTE Academic and IELTS Academic. London: Pearson.
- Council of Europe / Language Policy Unit (Strasbourg)-Project LIAM:www.coe.int/lang-migrants.
- Council of Europe (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.
- Dorans, N.J., Lyu, C.F., Pommerich, M., & Houston, W.M. (1997). Concordance between ACT Assessment and recentered SAT I sum scores. *College & University*, 73, 24-34.
- Ecctis (2023). *Referencing the LANGUAGECERT General and Academic tests to the Common European Framework of Reference for Language (CEFR)*. London: Ecctis.
- Elliot, M., Blackhurst, A., O'Sullivan, B., Clark, T., Dunlea, J., & Saville, N. (2021). Aligning IELTS and PTE-Academic: A measurement study. In N. Saville, B. O'Sullivan & T. Clark (Eds.), *IELTS Partnership Research Papers: Studies in Test Comparability Series*, No. 2, (pp. 42-64). IELTS Partners: British Council, Cambridge Assessment English and IDP: IELTS Australia
- ETS. (2010). *Linking TOEFL iBT scores to IELTS scores – a research report*. <https://www.ets.org/pdfs/toefl/linking-toefl-ibt-scores-to-ielts-scores.pdf>
- Green, A., Ünalı, A., & Weir, C. (2010). Empiricism versus connoisseurship: Establishing the appropriacy of texts in tests of academic reading. *Language Testing*, 27(2), 191-211.
<https://doi.org/10.1177/0265532209349471>

- Hatch, E. M., & Lazaraton, A. (1991). *The research manual: Design and statistics for applied linguistics*. New York, NY: Newbury House Publishers.
- House, G. (2010). Postgraduate Education in the UK. HEPI Analytical Report 1. *Higher Education Policy Institute*.
- Isaacs, T., Trofimovich, P., Yu, G., & Chereau, B. M. (2015). Examining the linguistic aspects of speech that most efficiently discriminate between upper levels of the revised IELTS pronunciation scale. *IELTS research reports online series*, 4.
- Knoch, U. (2021). A guide to English language policy making in higher education. International Education Association of Australia (IEAA). www.ieaa.org.au.
- Knoch, U., & Fan, J. (2024). Test score comparison tables: How well are they serving test users? *Language Testing*, 41(3). <https://doi.org/10.1177/02655322241239348>
- Koizumi, R., Agawa, T., Asano, K. et al. Skill profiles of Japanese English learners and reasons for uneven patterns. *Lang Test Asia* 12, 53 (2022). <https://doi.org/10.1186/s40468-022-00203-3>
- LaFlair G. T., Settles B. (2019). *Duolingo English Test: Technical manual* (Duolingo Research Report). <https://s3.amazonaws.com/duolingo-papers/other/Duolingo%20English%20Test%20-%20Technical%20Manual%202019.pdf>
- Lam, R. (2021). Using eportfolios to synergise assessment of, for, as learning in EFL writing. *European Journal of Applied Linguistics and TEFL*, 10(1), 101-120.
- Lampropoulou, L. (2022). Interactional competence and the role role play plays. *International Journal of TESOL Studies*, 4(1), 32-47. doi.org/10.46451/ijts.2022.01.04. · Mar 30, 2022
- Mayordomo, R. M., Espasa, A., Guasch, T., & Martínez-Melo, M. (2022). Perception of online feedback and its impact on cognitive and emotional engagement with feedback. *Education and Information Technologies*, 27(6), 7947-7971.
- Mežek, Š., McGrath, L., Negretti, R., & Berggren, J. (2022). Scaffolding L2 academic reading and self-regulation through task and feedback. *TESOL Quarterly*, 56(1), 41-67.
- North, B. (2021). An introduction to the theme 'Uneven profiles' [Language Assessment for Migrants' Integration (LAMI) Workshop]. [https:// www. youtube. com/ watch?v= oh9or yjv4f k&t= 1s](https://www.youtube.com/watch?v=oh9or yjv4f k&t= 1s).

- Parkinson, M., Howe, J., & Azarias, J. (2023). *Review of the migration system final report*. Australian Government Department of Home Affairs. Retrieved from https://parlinfo.aph.gov.au/parlInfo/download/media/pressrel/9137888/upload_binary/9137888.pdf;fileType=application%2Fpdf
- Pommerich, M. & Dorans, N. J. (Eds.) (2004). *Concordance* [Special issue]. *Applied Psychological Measurement*, 28(4), 216–289.
- Saville, N., O'Sullivan, B., & Clark, T. (Eds.) (2021). Investigating the relationship between IELTS and PTE-Academic. *IELTS Partnership Research Papers: Studies in Test Comparability Series*, No. 2. IELTS Partners: British Council, Cambridge Assessment English and IDP: IELTS Australia.
- Taylor, L., & Chan, S. (2015). *IELTS Equivalence Research Project* (GMC 133).

Appendix A: Exam interval analysis

To investigate how the length of time between the two exams might have affected the results, the data was split between test takers who had completed both within the targeted 90-day window ($n = 868$), and those who had completed them farther apart than 90 days ($n = 140$). This created four groups of test takers: those who took the tests within a 90-day period and took LanguageCert Academic first (Group 1) or took IELTS Academic first (Group 2); those who took the tests farther apart than 90 days and took LanguageCert Academic first (Group 3) or took IELTS Academic first (Group 4). Results were analysed descriptively and inferentially. Table 14 and Table 15 offer an initial summary of the differences between Groups, with Table 14 featuring exams taken within a 90-day period (Groups 1 and 2) and Table 15 featuring exams which were taken more 90 days apart (Groups 3 and 4).

Table 29. *LanguageCert Academic and IELTS Academic overall score comparison: Duration between exams within a 90-day period*

Group		First exam taken	LanguageCert Academic Score		IELTS Academic Score	
	<i>n</i>		<i>Mean</i>	<i>St. Dev.</i>	<i>Mean</i>	<i>St. Dev.</i>
1	466	LanguageCert Academic	64.35	12.89	6.34	.92
2	402	IELTS Academic	64.43	11.81	6.26	.84

Table 30. *LanguageCert Academic and IELTS Academic overall score comparison: Duration between exams > 90 days*

Group		First exam taken	LanguageCert Academic Score		IELTS Academic Score	
	<i>n</i>		<i>Mean</i>	<i>St. Dev.</i>	<i>Mean</i>	<i>St. Dev.</i>
3	83	LanguageCert Academic	56.53	10.41	5.84	.75
4	57	IELTS Academic	61.93	12.94	6.17	.95

Comparing Table 15 with Table 14, test takers who completed the two tests farther apart than 90 days, (Groups 3 and 4), scored lower on average (58.73 on LanguageCert, 5.97 on IELTS) than those taking the tests within a 90-day period (Groups 1 and 2) (64.39 on LanguageCert, 6.30 on IELTS). This was consistent across all comparisons. Table 14 and Table 15 show that when test takers performed well on the first exam taken, they tended to perform well on the second, and when they performed relatively poorly on the first exam taken, they also performed relatively poorly on the second.

To facilitate further investigation using inferential statistics, scores on the two exams were converted to the same scale. Each of the 1,008 LanguageCert Overall scores were converted

using the equivalencies from Table 22. Mean scores and standard deviations were re-calculated on this basis, followed by dependent samples *t*-tests.

Group 1 included 466 test takers. When converted to the IELTS band scale, their LanguageCert Academic scores displayed a mean of 6.30 (SD = .84) compared to the IELTS Academic mean of 6.34 (SD = .92) shown in Table 14. A dependent samples *t*-test showed the difference was not significant ($p > .05$).

For the 402 test takers in Group 2, the converted LanguageCert Academic scores displayed a mean of 6.30 (SD = .76) compared to the IELTS Academic mean of 6.26 (SD = .84) shown in Table 14. Like Group 1, a dependent samples *t*-test showed the difference was not significant ($p > .05$).

In Group 3, 83 test takers took the two exams farther than 90 days apart and took LanguageCert Academic first. When LanguageCert Academic scores were converted to IELTS bands, this Group displayed a mean of 5.83 (SD = .64) compared to the IELTS Academic mean of 5.84 (SD = .75) shown in Table 15. A dependent samples *t*-test showed the difference was not significant ($p > .05$).

In Group 4, 57 test takers took the two tests farther apart than 90 days and took IELTS Academic first. When LanguageCert Academic scores were converted to IELTS bands, this Group displayed a mean of 6.11 (SD = .78) compared to their IELTS Academic mean score of 6.17 (SD = .95) shown in Table 15. Commensurate with the other differences reported in this section, a dependent samples *t*-test showed the difference was not significant ($p > .05$).

The above results support the notion that for the present study, test takers performed similarly regardless of whether they took the exam within 90 days or at an interval greater than 90 days. This result - in conjunction with the counterbalanced study design - justifies the inclusion of data from test takers who took the tests more than 90 days apart.

Appendix B: Concordance study – Candidate Questionnaire

Candidate's First Name:	
Candidate's Last Name:	

The following questionnaire forms part of a research project conducted by LANGUAGECERT. Please take time to read the following information carefully. **This study aims to explore how candidates perform in the LANGUAGECERT Academic and the LANGUAGECERT General exams.**

Your personal and exam data will be anonymised and encrypted on password-protected computers. Hard copies of any data will be stored securely. All data will be held within our storage systems in line with LANGUAGECERT's policy for data security. Anonymised data will also be shared with the Centre for Research in English Language Learning and Assessment. Your data will be kept securely for a minimum of ten years.

- ☐ I confirm that I have read and I understand the information for the study.
- ☐ I understand that any information given by me may be used in a research study and it may also be used in publications or presentations by the researchers, but my personal information will not be included, and I will not be identifiable.
- ☐ I understand that my name will not appear in any reports, articles or presentations.
- ☐ I agree to take part in the above study.
- ☐ I have received and studied the familiarisation material for LANGUAGECERT and IELTS exams.

Background details		
1	I am years old	<input type="checkbox"/> <18 <input type="checkbox"/> 19-23 <input type="checkbox"/> 24-30 <input type="checkbox"/> 31-40 <input type="checkbox"/> 41 or above
2	I am	<input type="checkbox"/> Male <input type="checkbox"/> Female <input type="checkbox"/> Non – Binary <input type="checkbox"/> Prefer not to say
3	I live in (country)	<input type="checkbox"/> Greece <input type="checkbox"/> China <input type="checkbox"/> India <input type="checkbox"/> Australia <input type="checkbox"/> Other:
4	My first language is
5	My highest education level is ...	<input type="checkbox"/> Primary (up to age 12) <input type="checkbox"/> Secondary (up to age 18) <input type="checkbox"/> Postsecondary (example, Bachelor's degree) <input type="checkbox"/> Higher Degree (example, Master's degree or Doctorate)
6	I have taken one of these English language tests before...	<input type="checkbox"/> TOEFL iBT Date: (MM/YY), Score: <input type="checkbox"/> Pearson Test of English Date: (MM/YY), Score:

LANGUAGECERT test details				
7	The LANGUAGECERT test I am taking is....	<input type="checkbox"/> Academic	<input type="checkbox"/> General	<input type="checkbox"/> Not sure
8	For me, the LANGUAGECERT test format (the number and type of questions on the test) is...	<input type="checkbox"/> Very familiar <input type="checkbox"/> Familiar <input type="checkbox"/> A little familiar <input type="checkbox"/> Not familiar		
9	I have taken a LANGUAGECERT test before:	<input type="checkbox"/> Yes / Date: (MM/YY), Score: <input type="checkbox"/> No		
10	I plan to use my LANGUAGECERT test results for... [you may choose more than one answer]	<input type="checkbox"/> a higher education extended course (3 months or more) <input type="checkbox"/> a higher education short course (3 months or less) <input type="checkbox"/> other education purposes <input type="checkbox"/> registration as a healthcare professional <input type="checkbox"/> professional registration (NOT healthcare) <input type="checkbox"/> employment purposes <input type="checkbox"/> immigration purposes <input type="checkbox"/> personal reasons <input type="checkbox"/> other purpose:		

IELTS test details				
11	The IELTS test I am taking is....	<input type="checkbox"/> Academic	<input type="checkbox"/> General	<input type="checkbox"/> Not sure
12	For me, the IELTS test format (the number and type of questions on the test) is...	<input type="checkbox"/> Very familiar <input type="checkbox"/> Familiar <input type="checkbox"/> A little familiar <input type="checkbox"/> Not familiar		
13	I have taken an IELTS test before:	<input type="checkbox"/> Yes / Date: (MM/YY), Score: <input type="checkbox"/> No		
14	I plan to use my IELTS test results for... [you may choose more than one answer]	<input type="checkbox"/> a higher education extended course (3 months or more) <input type="checkbox"/> a higher education short course (3 months or less) <input type="checkbox"/> other education purposes <input type="checkbox"/> registration as a healthcare professional <input type="checkbox"/> professional registration (NOT healthcare) <input type="checkbox"/> employment purposes <input type="checkbox"/> immigration purposes <input type="checkbox"/> personal reasons <input type="checkbox"/> other purpose:		



LANGUAGECERT is an Awarding Organisation recognised by Ofqual. It spearheads innovations in language assessment and certification, providing high-quality services to the global learners' community. It is a UK-based member of the PeopleCert Group, a global leader in the certification industry, that delivers millions of exams in over 200 countries.

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

languagecert.org | info@languagecert.org