

The Oxford ELLT: Evolving Test Design for Modern Academic Contexts

June 2026

Author: Kamila Dolejsi

Introduction

Language tests rarely stay still for long. As teaching, technology, and student needs evolve, so must the ways we measure language ability. The Oxford English Language Level Test (ELLT) is one such example: a relatively new assessment that has grown quickly in both scale and recognition, while continuing to adapt to changing academic and global contexts. This paper provides an overview of the ELLT—its structure, development, and recent revisions—showing how it aims to balance flexibility, reliability, and authenticity in modern language assessment.

The Oxford ELLT: Structure and Delivery

The Oxford English Language Level Test (ELLT) is a modular, four-skills online assessment that evaluates reading, listening, writing, and speaking. Delivered on demand, the test allows candidates to schedule their assessment at a time that suits their individual circumstances, rather than adhering to fixed dates or locations. This flexibility supports accessibility, particularly for candidates who may face logistical or technological constraints. As Xi (2010) notes, technology-mediated assessment has the potential to broaden access for diverse test-taking populations.

Candidates may begin with either the listening or reading module. After completing both receptive skills, they proceed to writing and then speaking. The modular design allows candidates to complete all components in one sitting or spread them across multiple sessions, depending on their circumstances. This approach reflects principles of inclusive assessment design, with Kunnan (2018) emphasising that fairness involves enabling all test-takers to demonstrate their abilities under appropriate conditions.

Scoring methods vary by skill: reading and listening are automatically marked, while writing and speaking are assessed by trained and standardised examiners. The productive skills are also integrated; for example, speaking includes follow-up questions related to the candidate's writing task. This design supports a more holistic evaluation of language ability, aligning with Weir's (2005) argument that valid assessments should reflect authentic language use and the interaction between skills. Test results are typically issued within 48 hours and shared directly with partner institutions.

Background

Introduced in 2020 in response to disruptions caused by the COVID-19 pandemic, the Oxford ELLT initially supported Oxford International embedded colleges when traditional testing routes were unavailable. Since then, it has been taken by approximately 200,000 candidates across more than 150 locations, including individuals from at-risk contexts. It is now accepted by around 65% of UK universities, reflecting its growing recognition within the higher education sector.

Oxford ELLT in Numbers

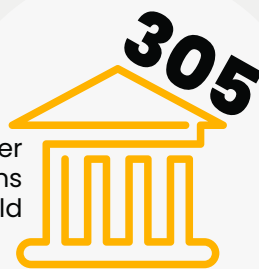


The Oxford ELLT has been in existence for 6 years



215,000

Over 215,000 total number of tests have been taken to date



We partner with 305 higher education institutions around the world



151 nationalities tested to date

350+

Over 350 at-risk university applicants tested to date



Content Development and Quality Assurance

The Oxford ELLT is supported by a large and continuously expanding item bank, currently comprising approximately 80 speaking tests, 80 writing tests, 120 reading tests, and 135 listening tests. This enables the creation of multiple parallel forms and enhances test security through randomisation.

All materials are developed in-house by experienced ELT professionals and updated on a bimonthly basis. Topics are selected to balance authenticity, accessibility, and cultural sensitivity, covering areas such as technology, education, health, and the environment.

Item development follows clearly defined principles and is grounded in internationally recognised proficiency frameworks. Central to this is the Common European Framework of Reference for Languages (CEFR), developed by the Council of Europe to provide a standardised basis for describing language ability. The CEFR organises proficiency into six levels (A1–C2), grouped into Basic, Independent, and Proficient users, and defines each level through detailed “can do” descriptors. These descriptors emphasise observable performance across listening, reading, speaking, and writing, shifting the focus from abstract grammatical knowledge to practical communicative competence.

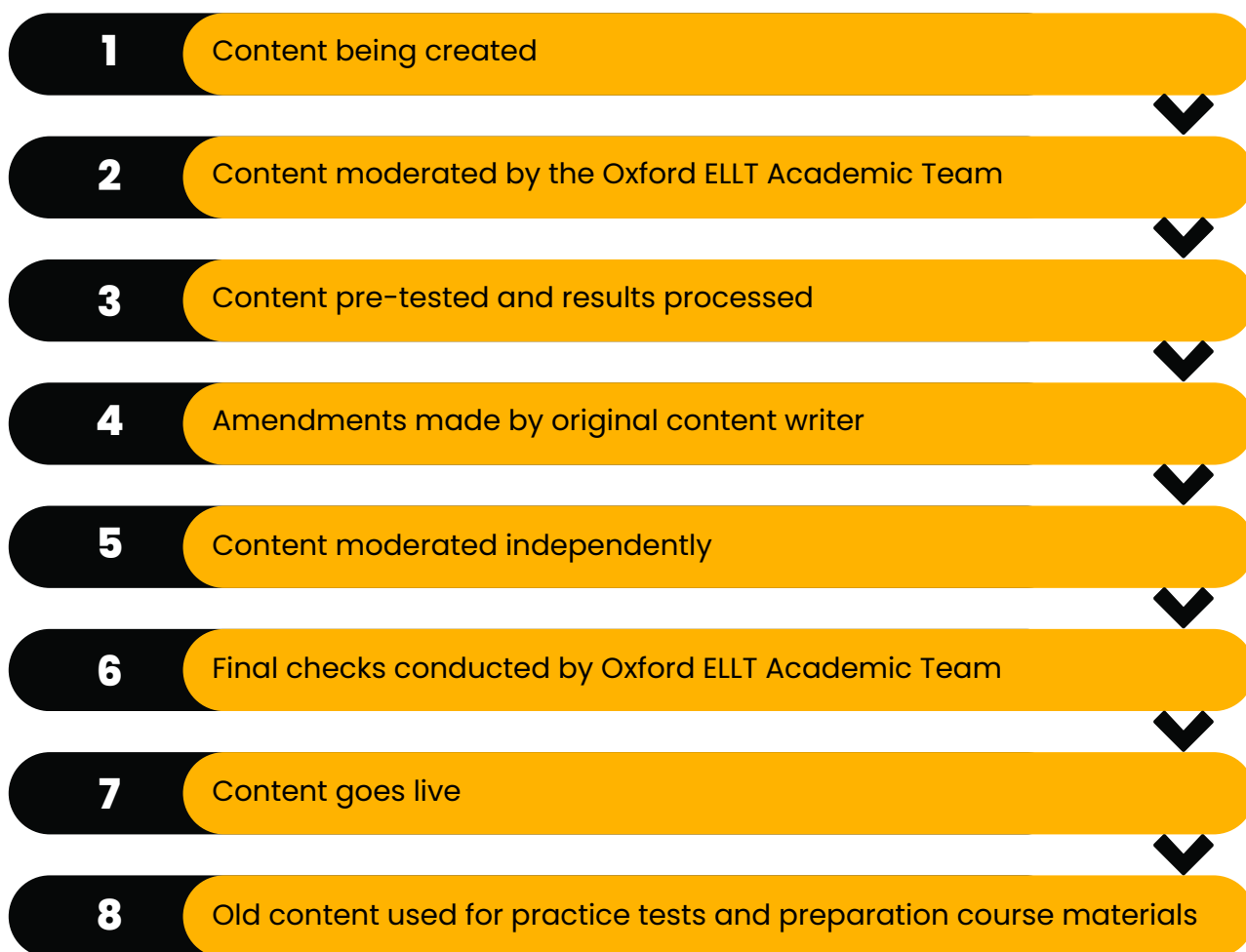
To complement the CEFR, the Oxford ELLT also draws on the EAQUALS Core Inventory for General English (developed in collaboration with the British Council), which operationalises CEFR levels by specifying the grammatical structures, lexical ranges, and functional language associated with each level. This framework provides detailed guidance on the linguistic resources learners need in order to perform the communicative tasks described in CEFR descriptors. In particular, it highlights the relationship between form and function, outlining how learners progress from basic sentence patterns and high-frequency vocabulary at A levels, through expanded structural and lexical range at B levels, to nuanced, flexible, and idiomatic language use at C levels.

Texts and tasks are analysed using levelling tools and aligned with these frameworks to ensure consistency, validity, and appropriate difficulty. All materials undergo multiple drafts and are subject to strict referencing requirements.

Quality assurance is maintained through a multi-stage moderation process, including both internal and external review. Materials are evaluated for construct validity, level appropriateness, and originality, and are developed within an eight-week production cycle. Content older than three months is retired from active use and repurposed for practice materials.

A central component of the validation process is pre-testing under authentic conditions. Test materials are trialled with adult learners in the UK, US, and Canada, using cohorts with established proficiency levels. Each set of items is tested across at least two CEFR levels to ensure accurate calibration and alignment with proficiency standards.

Following revision, items undergo further moderation before final approval. Only those meeting all criteria proceed to publication on the testing platform.



Recent Revisions and Design Principles

The 2026 revisions represent a coordinated effort to enhance the test's validity and authenticity while maintaining its established structure. These changes are informed by empirical research, competitor analysis, and consultation with higher education stakeholders. In line with Fulcher (2010), the revisions aim to better align the test with target language use in academic contexts, while also reflecting Bachman and Palmer's (1996) principle that test usefulness depends on balancing innovation with continuity.

Reading Component

The reading component has been revised to include three texts of varying lengths, replacing the previous two-text format. While the overall test duration remains 40 minutes, candidates now engage with a short text (approximately 200 words), a medium-length text (around 400 words), and a longer text (500–600 words). The total number of questions has been slightly reduced from 32 to 30, while maintaining the same total score through weighted items.

Task types have been diversified to reflect a broader range of reading subskills. The first text (B1/B1+) focuses on paragraph sequencing, the second (B2) on gap-fill completion at phrase level, and the third (C1) on multiple-choice comprehension. Question difficulty spans A2 to C1 levels. Across these tasks, candidates are required to demonstrate a range of reading abilities, including identifying main ideas and supporting detail, understanding specific factual information, recognising cohesive devices and discourse markers, and interpreting implied meaning and author viewpoint. They must also distinguish between stated and non-stated information and engage with vocabulary in context. This range aligns with Grabe's (2009) view that reading proficiency involves multiple interacting processes, including inference and discourse organisation.

Listening Component

The listening component retains its overall structure: a 25-minute test with three audio inputs and 32 questions. However, the second task has been redesigned, replacing multiple-choice questions with a note-completion format requiring word- or phrase-level responses.

The listening inputs include a monologue played once, a second recording (monologue or dialogue) played twice, and a multi-speaker analysis task played twice. Task levels range from B1 to C1, with complexity spanning A2 to C1. Across the component, candidates are expected to follow extended spoken discourse, identify key points and supporting information, and distinguish between gist, key detail, and specific information. They must also interpret speaker purpose, tone, and viewpoint, distinguish between similar opinions, and transfer information accurately. The introduction of note completion reflects Flowerdew's (1994) observation that academic listening often involves processing and recording information in real time.

Writing Component

The writing component now includes two tasks, increasing the total duration from 45 to 50 minutes. The first is an integrated task requiring candidates to summarise a short reading passage (80–100 words based on a 170–200-word text). The second remains an essay response, with a slightly adjusted length of 190–250 words. This addition reflects Weigle's (2004) argument that integrated tasks better represent authentic academic writing practices, while retaining the essay format aligns with Hyland's (2003) emphasis on extended writing in academic discourse. These two tasks are not equally weighted; task one is assessed with a global mark which carries 30% of the overall grade. Task two is assessed using writing rubrics and carries 70% of the overall grade for the writing component.

These tasks assess a range of writing abilities, including paraphrasing and clear expression of ideas, organisation and paragraph structure, and the ability to present and evaluate arguments. Candidates are also expected to demonstrate range and control of vocabulary, grammatical accuracy and complexity, and the effective use of examples to support reasoning.

Speaking Component

The speaking component retains its overall duration of 20–25 minutes but introduces a key procedural change. The previously required prepared presentation has been removed and replaced with a short, prompt-based monologue. Candidates are given up to three prompts of increasing complexity and 45 seconds to prepare before speaking for 2–3 minutes.

Other elements remain unchanged, including an examiner-led discussion based on the candidate's writing and a further discussion based on a visual prompt. The component assesses candidates' ability to produce extended spoken responses, organise and clarify ideas, and respond spontaneously in real time. It also evaluates their ability to justify viewpoints, engage in interactive communication, respond flexibly to follow-up questions, and maintain fluency, coherence, and clear pronunciation. This shift prioritises spontaneous communication, consistent with Luoma's (2004) view that speaking assessments should elicit real-time language use. All tasks are equally weighted; while tasks 2 and 3 are assessed using speaking rubrics, part 3 allows the examiner to move the prompts up and / or down the levels and is therefore awarded a global mark.

Conclusion

The Oxford ELLT continues to evolve while maintaining its core structure and purpose. The 2026 revisions strengthen its alignment with real-world academic language use by expanding task types, integrating skills, and emphasising authentic communication. Supported by rigorous development and validation processes, the test aims to provide a reliable, flexible, and accessible measure of English proficiency for contemporary learners.

Oxford ELLT Assessment criteria for Writing

Task 1		Task 2			
TASK 1 Global Score	OELLT Level	Task Achievement	Organisation	Accuracy of lexis and grammar	Range of lexis and grammar
The summary demonstrates very limited understanding of the source text. Important information is often omitted, copied directly, or presented inaccurately. The response may be incomplete, unclear, or largely irrelevant to the task.	ELLT 0-1	The writing did not relate to the assigned topic. It did not follow the conventions of an essay-style text and it was significantly shorter or longer. The content was predominantly irrelevant.	The writing was not well organised. Basic cohesive devices such as "and", "because" or "then" were used sporadically or inaccurately. The reader could not follow due to lack of organisation.	The grammar and lexis were basic and limited in accuracy with some simple structures and lexis used correctly. Errors were very frequent and caused strain on the reader.	The range was limited to very basic sentence patterns and grammatical forms that are inappropriate to the task and used repetitively. The candidate demonstrated sufficient vocabulary to clearly express a straightforward everyday point.
The summary demonstrates a basic understanding of the main ideas in the source text. Some key information is identified, but important points may be missed, repeated, or included without clear organisation. The candidate relies heavily on the language of the original text and shows limited ability to summarise information in their own words.	ELLT 2-4	The writing partially met the task. Some content presented was irrelevant. Some awareness of essay-style conventions is shown. Can produce a simple, straightforward text that links short statements and ideas together that links a series of statements and ideas together to list advantages and disadvantages or give an opinion.	The writing was not particularly well-organised, but there was some attempt to link ideas. A range of cohesive devices were used, though often inaccurately. Used paragraphing, but not always logically.	Simple structures and lexis were largely accurate, but basic mistakes still appear systematically. Attempted to use some complex structures, but with limited accuracy. Errors in spelling, word formation, lexis and complex structures could potentially cause some strain on the reader, but writing is generally intelligible.	The range of basic grammar structures and vocabulary demonstrated was good but limited to everyday language and at times repetitive, using the same simple language to list advantages and disadvantages, for example. Some awareness of complex language was shown, although more advanced lexis was often used inappropriately.
The summary demonstrates a good understanding of the main ideas and supporting information in the source text. Key points are selected appropriately and organised logically. The candidate is able to summarise information in their own words, although paraphrasing and organisation may not always be fully effective.	ELLT 5-7	The writing met the task. The main ideas were relevant, but are often unclear or inadequately supported. A awareness of formal essay conventions is shown, despite occasional use of informal language. The main ideas and arguments in the text are detailed and clear.	The writing was logically organised and coherent using a variety of linking words and cohesive devices to link and contrast ideas mostly accurately. Though some statements may be misplaced, the main arguments are clearly distinguishable through appropriate use of paragraphing.	A very good degree of accuracy and control over simple structures and lexis, and some complex language. Spelling, word formation and punctuation are reasonably accurate, but may show errors and signs of mother tongue influence. Shows a relatively high degree of grammatical and lexical control, and mistakes which lead to misunderstanding are rare.	The range of grammar structures and vocabulary demonstrated was good. Everyday as well as more advanced or topic-specific vocabulary was used appropriately to develop the argument and to vary formulation to avoid frequent repetition. Some lexical gaps or inaccuracies occasionally cause confusion for the reader.
The summary demonstrates a good understanding of the main ideas and supporting information in the source text. Key points are selected appropriately and organised logically. The candidate is able to summarise information in their own words, although paraphrasing and organisation may not always be fully effective.	ELLT 8-10	The writing was well-structured, covered more complex aspects of the topic, and fully met the task. All content was relevant to the task, and the main ideas were well supported and developed. The writing addresses all parts of the task and for the most part presents a clear position throughout. A good awareness of register and formal writing conventions is shown, despite rare resort to informal language.	The writing was well organised and showed controlled use of organisational patterns, connectors and cohesive devices. A wide variety of cohesive devices and organisational patterns were used with flexibility. Paragraphs have a clear central topic. Ideas are linked and/or contrasted within and between paragraphs. Layout and paragraphing are consistent and helpful.	A high degree of accuracy and flexibility of complex grammatical structures and lexis used. Slips or non-systematic errors and minor flaws in spelling, sentence structure or punctuation still occur, but they are rare. Errors were rare, difficult to spot and minor, occurring with high-level language or as slips	The range of complex grammar structures and vocabulary demonstrated was great. Uses a wide range of complex and advanced grammatical forms and high-level lexis with flexibility. Incorporates a range of language features such as idioms and collocations, though uses may not always be appropriate. Can use a broad range of complex grammatical structures appropriately and with considerable flexibility.
The summary demonstrates a strong understanding of the source text and the relationships between ideas. Information is selected, organised, and condensed effectively, focusing on the most important points. The candidate paraphrases successfully and demonstrates clear control of summarising as a mediation task.	ELLT 11-12	The writing fully met the task. The genre and length were met. The reader was fully informed, all ideas were fully extended and supported. The writing addresses all parts of the task, presents a clear position throughout and a personal reflection was evident even with a conceptually difficult subject. Excellent awareness of formal writing conventions and academic register is shown. The candidate clearly and effectively distinguishes their own ideas and opinions from others to support their own argument.	The writing was logically organised and coherent. The text was clear, smoothly flowing, well-structured, showing controlled use of organisational patterns, connectors and cohesive devices. The text provided an appropriate and effective logical structure which helps the reader identify significant points, with ideas and arguments linked effectively to develop a fairly complex or nuanced argument.	A very high degree of flexibility and accuracy of the grammar structures and lexis used. The candidate maintains consistent grammatical control of complex language. Consistently correct and appropriate use of vocabulary. Punctuation is used consistently well. Writing is orthographically free of error.	The range of complex grammar structures and vocabulary demonstrated was exceptional. Can effectively express and exploit a very wide range of language to formulate thoughts precisely, give emphasis, differentiate and eliminate ambiguity. Smooth and flowing incorporation of language features, with an excellent command of a broad lexical repertoire and strong awareness of connotative levels of meaning.

Oxford ELLT Assessment criteria for Speaking

OELLT Level	Task Achievement	Accuracy and appropriacy of language	Fluency and Discourse Management	Pronunciation
ELLT 0 – 1	The candidate is unable to complete the task successfully. Responses are extremely limited and may be memorised or unrelated to the topic. Communication is restricted to simple words, phrases, or very short responses.	The lexis and grammar used were very basic and largely inappropriate for the task. Errors were very frequent and caused a break-down in communication. A few basic sentences were error free. Candidate could handle very short social exchanges but was rarely able to understand enough to keep conversation going of their own accord.	Content was obviously read from a prepared script or presented it in an over rehearsed	Pronunciation consists of isolated sounds and is largely unintelligible. Communication cannot be achieved. Cannot be understood in an international context.
ELLT 2 – 4	The candidate partially completes the task. Some parts of the task may be missing, inadequately addressed, or only minimally developed. Responses are generally simple and may require support from the examiner. Ideas are expressed but are often limited in detail or explanation.	The lexis and grammar used were basic, but appropriate for the task. Used simple structures, but with limited control. Errors were frequent and at times impeded communication, but candidate could handle very simple exchanges on a number of familiar topics.	Pronunciation consists of isolated sounds and is largely unintelligible. Communication cannot be achieved. Cannot be understood in an international context.	Delivery may pose some strain on the listener. The message is conveyed in an intelligible way and is clear despite regular mispronunciation of individual sounds and words they are less familiar with. Intonation and stress are frequently influenced by accent.
ELLT 5–7	The candidate completes the task and addresses all parts. Ideas are developed with relevant supporting information and responses are generally clear, organised, and easy to follow. The candidate can explain opinions and provide reasons and examples to support their views.	The lexis and grammar selected were suitable. Full control of simple structures and lexis use. Attempted complex lexis/ structures but with limited flexibility. Use of complex structures with errors which did not impede communication.	The delivery was natural. Some signs of hesitation and repetition without impeding the communication, sometimes communicated spontaneously. Ideas were linked using simple and complex linking devices and/or signposting language. The candidate was able to adjust what they say and the means of expressing it to the situation and the recipient and adopt a level of formality appropriate to the circumstances.	Delivery is not completely natural. Sometimes does not use appropriate intonation or stress but articulates individual sounds clearly. The overall message remains clear and is intelligible throughout. It is generally understood in an international academic setting without too much difficulty.
ELLT 8–10	The candidate fully completes the task. Ideas are well developed and supported through explanation, examples, comparison, and justification. Responses are detailed, coherent, and demonstrate the ability to discuss topics in depth and from different perspectives.	The lexis and grammar selected were suitable for the subject matter. Simple and complex lexis/structures were used with few errors. Sentences consist of mainly complex lexis/structures. The language was largely error-free with a high degree of grammatical accuracy and self-correction generally occurs.	The delivery was natural. Ease of expression in even longer complex stretches of speech. Ideas were effectively linked using a wide range of linking devices and/or signposting language. The candidate was able to modify their expression to express degrees of commitment or hesitation, confidence or uncertainty. The candidate was able to express themselves spontaneously, fluently and effortlessly.	Delivery is natural. Produces smooth intelligible spoken discourse with only occasional lapses in control of stress, rhythm and/or intonation, which do not affect intelligibility or effectiveness. Pronunciation does not cause communication problems in any international context.
ELLT 11–12	The candidate fully completes the task to a very high standard. Complex, abstract, and hypothetical ideas are discussed confidently and effectively. Responses demonstrate sophisticated reasoning, excellent development of ideas, and the ability to communicate nuanced viewpoints with ease.	The grammar and lexis selected were suitable for the subject matter and showed a wide range of complexity. The language was mostly error-free and self-corrected.. Showed a good command of idiomatic expressions and colloquialisms. Was able to backtrack and restructure around a difficulty so smoothly that the interlocutor is hardly aware of it.	The delivery was naturally fluent and without any hesitations or repetition. Ideas were effectively linked using an exceptional range of linking devices and/or signposting language. Can converse comfortably and appropriately, unhampered by any linguistic limitations.	Pronunciation is clearly understood by both L1 and L2 English speakers. Was able to employ the full range of phonological features in the target language with a high level of control so that the finer points of their message are clear and precise. . Pronunciation allows for effortless and nuanced communication across a wide range of topics, contexts, and social situations.

References

Bachman, L.F. and Palmer, A.S. (1996) *Language testing in practice*. Oxford: Oxford University Press.

Council of Europe (2020) *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume*. Strasbourg: Council of Europe Publishing.

Flowerdew, J. (1994) *Academic listening: Research perspectives*. Cambridge: Cambridge University Press.

Fulcher, G. (2010) *Practical language testing*. London: Hodder Education.

Grabe, W. (2009) *Reading in a second language: Moving from theory to practice*. Cambridge: Cambridge University Press.

Hyland, K. (2003) *Second language writing*. Cambridge: Cambridge University Press.

Kunnan, A.J. (2018) *Evaluating language assessments*. New York: Routledge.

Luoma, S. (2004) *Assessing speaking*. Cambridge: Cambridge University Press.

Weigle, S.C. (2004) 'Integrating reading and writing in a competency test for non-native speakers of English', *TESOL Quarterly*, 38(2).

Weir, C.J. (2005) *Language testing and validation: An evidence-based approach*. Basingstoke: Palgrave Macmillan.

Xi, X. (2010) 'Automated scoring and feedback systems in language testing', *Language Testing*, 27(3).

The Oxford ELLT Reflections on test revision June 2026

Oxford International Digital Institute
259 Greenwich High Road
Greenwich
London
UK
SE10 8NB



oxfordellt.com