TABELLA RICONOSCIMENTO CERTIFICAZIONI INTERNAZIONALI – SETTEMBRE 2024

Livello minimo richiesto per il riconoscimento dell'esame	Quadro Europeo B1	Quadro Europeo B2	Quadro Europeo C1
Inglese – Cambridge English Qualifications (General English, Business English, Schools)	**CAMBRIDGE ENGLISH SCALE [1]** da 140 a 152 = 28 da 153 a 159 = 29 da 160 a 170 = 30 più di 170 = 30 e lode	**CAMBRIDGE ENGLISH SCALE [1]** da 160 a 172 = 28 da 173 a 179 = 29 da 180 a 190 = 30 più di 190 = 30 e lode	**CAMBRIDGE ENGLISH SCALE [1]** da 180 a 192 = 28 da 193 a 199 = 29 da 200 a 205 = 30 più di 205 = 30 e lode
Inglese – Cambridge IGCSE "First Language English" (E1L) Syllabus: 0500, 0990. **Per il riconoscimento, devono essere valutate tutte le quattro abilità.**	**Consultare le tabelle di equivalenza [2]** Tutte le abilità B1 = 28 Tutte le abilità superiori o uguali a B2 = 30 e lode Altre combinazioni, con tutte le abilità superiori o uguali a B1 (Es. B1 + B1 + B2) = 29	**Consultare le tabelle di equivalenza [2]** Tutte le abilità B2 = 28 Tutte le abilità superiori o uguali a C1 = 30 e lode Altre combinazioni, con tutte le abilità superiori o uguali a B2 = 29	**Consultare le tabelle di equivalenza [2]** Tutte le abilità C1 = 28 Altre combinazioni, con tutte le abilità superiori o uguali a C1 = 30
Inglese – Cambridge IGCSE "English as a Second Language" (E2L) Syllabus: 0511, 0991, 0510, 0993. **Per il riconoscimento, devono essere valutate tutte le quattro abilità.**	**Consultare le tabelle di equivalenza [3]** Tutte le abilità B1 = 28 Tutte le abilità superiori o uguali a B2 = 30 e lode Altre combinazioni, con tutte le abilità superiori o uguali a B1 (Es. B1 + B1 + B1 + B2) = 29	**Consultare le tabelle di equivalenza [3]** Tutte le abilità B2 = 28 Altre combinazioni, con tutte le abilità superiori o uguali a B2 = 29	
Inglese – Trinity (ISE) INTEGRATED SKILLS IN ENGLISH	**ISE I** Pass = 28 Pass with merit = 29 Pass with distinction = 30 **ISE II, ISE III, ISE IV** = 30 e lode	**ISE II** Pass = 28 Pass with merit = 29 Pass with distinction = 30 **ISE III, ISE IV** = 30 e lode	**ISE III** Pass = 28 Pass with merit = 29 Pass with distinction = 30 **ISE IV** = 30 e lode
Inglese - TOEFL ETS (Educational Testing Service) *TOEFL iBT® Scores*, *TOEFL iBT® Home Edition*, *MyBest™ Scores*	**Score** da 44 a 52 = 28 da 53 a 60 = 29 da 61 a 68 = 30 più di 68 = 30 e lode	**Score** da 69 a 77 = 28 da 78 a 85 = 29 da 86 a 93 = 30 più di 93 = 30 e lode	**Score** da 94 a 100 = 28 da 100 a 105 = 29 da 105 a 110 = 30 più di 110 = 30 e lode
Inglese – IELTS (Academic o General)	**IELTS** 4.0 = 28 4.5 = 29 5.0 = 30 più di 5.0 = 30 e lode	**IELTS** 5.5 = 28 6.0 = 29 6.5 = 30 più di 6.5 = 30 e lode	**IELTS** 6.5 = 28 7 = 29 7.5 - 8 = 30 più di 8.0 = 30 e lode
Inglese – Pearson Test of English – Academic (PTE Academic)	**PTE Academic** da 43 a 48 = 28 da 49 a 53 = 29 da 54 a 58 = 30 più di 58 = 30 e lode	**PTE Academic** da 59 a 65 = 28 da 66 a 70 = 29 da 71 a 75 = 30 più di 75 = 30 e lode	**PTE Academic** da 76 a 78 = 28 da 79 a 81 = 29 da 81 a 84 = 30 più di 84 = 30 e lode
Inglese – Pearson English International Certificate (PTE General)	**PTE General** **Level 2 (Intermediate)** pass = 28 pass with merit = 29 pass with distinction = 30 **Level 3 (Upper Intermediate)** pass at lower level (B1) = 28 pass, pass with merit/distinction = 30 e lode **Level 4 (Advanced)** = 30 e lode **Level 5 (Proficient)** = 30 e lode	**PTE General** **Level 3 (Upper Intermediate)** pass = 28 pass with merit = 29 pass with distinction = 30 **Level 4 (Advanced)** pass at lower level (B2) = 28 pass, pass with merit/distinction = 30 e lode **Level 5 (Proficient)** = 30 e lode	**PTE General** **Level 4 (Advanced)** pass = 28 pass with merit = 29 pass with distinction = 30 **Level 5 (Proficient)** pass at lower level (C1) = 28 pass, pass with merit/distinction = 30 e lode
Inglese – Oxford Test of English e Oxford Test of English Advanced **Per il riconoscimento, devono essere conseguiti tutti i quattro moduli.**	**Standardized score** da 81 a 90 = 28 da 91 a 100 = 29 da 101 a 110 = 30 più di 110 = 30 e lode	**Standardized score** da 111 a 120 = 28 da 121 a 130 = 29 da 131 a 139 = 30 più di 139 = 30 e lode	**Standardized score** da 141 a 150 = 28 da 151 a 160 = 29 da 161 a 169 = 30 più di 169 = 30 e lode
Inglese – LanguageCert International ESOL (Listening, Reading, Writing) + (Speaking) **Per il riconoscimento, devono essere conseguiti tutti i quattro moduli. In caso di moduli conseguiti a livelli diversi, il modulo di livello superiore sarà considerato come un "High pass" del livello inferiore.**	**B1 Achiever** (Listening, Reading, Writing) + (Speaking) Entrambi "Pass" = 28 "Pass" + "High pass" = 29 Entrambi "High pass" = 30 **B2 Communicator** = 30 e lode **C1 Expert** = 30 e lode **C2 Mastery** = 30 e lode	**B2 Communicator** (Listening, Reading, Writing) + (Speaking) Entrambi "Pass" = 28 "Pass" + "High pass" = 29 Entrambi "High pass" = 30 **C1 Expert** = 30 e lode **C2 Mastery** = 30 e lode	**C1 Expert** (Listening, Reading, Writing) + (Speaking) Entrambi "Pass" = 28 "Pass" + "High pass" = 29 Entrambi "High pass" = 30 **C2 Mastery** = 30 e lode
Inglese – Gatehouse Awards	**GA Entry Level Certificate in ESOL International (Entry 3) (Classic B1)** Pass = 28 Pas with merit = 30 **Classic B2, Classic C1, Classic C2** = 30 e lode	**GA Level 1 Certificate in ESOL International (Classic B2)** Pass = 28 Pas with merit = 30 **Classic C1, Classic C2** = 30 e lode	**GA Level 2 Certificate in ESOL International (Classic C1)** Pass = 28 Pas with merit = 30 **Classic C2** = 30 e lode
Inglese – British Council Aptis ESOL	**Overall CEFR Grade** B1 = 28 B2 = 30 e lode C = 30 e lode	**Overall CEFR Grade** B2 = 28 C = 30 e lode	**Overall CEFR Grade** C = 28

Livello minimo richiesto per il riconoscimento dell'esame	Quadro Europeo B1	Quadro Europeo B2	Quadro Europeo C1
Francese – Alliance Française (prima di settembre 2005)	**DELF 1° COMPLETO** da 10 a 13 = 28 da 14 a 17 = 29 da 18 a 20 = 30 **DELF 2° =** 30 e lode **DALF =** 30 e lode	**DELF 2°** da 10 a 13 = 28 da 14 a 17 = 29 da 18 a 20 = 30 **DALF =** 30 e lode	
Francese - Alliance Française (valido da settembre 2005)	**DELF B1** da 50 a 64 =28 da 65 a 74 = 29 da 75 a 100 = 30 **DELF B2, DELF C1, DELF C2 =** 30 e lode	**DELF B2** da 50 a 64 =28 da 65 a 74 = 29 da 75 a 100 = 30 **DELF C1, DELF C2 =** 30 e lode	
Francese – CCI Paris Diplomes de Français Professionel (DFP)	**DFP B1** da 60 a 69 = 28 da 70 a 79 = 29 da 80 a 100 = 30 **DFP B2, DFP C1, DFP C2 =** 30 e lode	**DFP B2** da 60 a 69 = 28 da 70 a 79 = 29 da 80 a 100 = 30 **DFP C1, DFP C2 =** 30 e lode	
Francese – EsaBac doppio diploma	**EsaBac generale / EsaBac Techno** 30 e lode	**EsaBac generale / EsaBac Techno** 30 e lode	
Spagnolo – DELE (Istituto Cervantes)	**DELE B1 Nivel Inicial** da 60 a 72 = 28 da 73 a 86 = 29 da 87 a 100 = 30 **DELE B2 Nivel Intermedio DELE C1 Nivel Superior** 30 e lode	**DELE B2 Nivel Intermedio** da 60 a 72 = 28 da 73 a 86 = 29 da 87 a 100 = 30 **DELE C1 Nivel Superior** 30 e lode	
Tedesco (Goethe-Institut)	**Goethe-Zertifikat B1 Erwachsene/Jugendliche** meno di 60 punti = NON SUPERATO da 60 a 69 = ausreichend = 28 da 70 a 79 = befriedigend = 29 da 80 a 89 = gut = 30 da 90 a 100 = sehr gut = 30 e lode **Goethe-Zertifikat B2 =** 30 e lode **Goethe-Zertifikat C1 =** 30 e lode **Goethe-Zertifikat C2 =** 30 e lode	**Goethe-Zertifikat B2 Erwachsene/Jugendliche** meno di 60 punti = NON SUPERATO da 60 a 69.5 = 28 da 70 a 79.5 = 29 da 80 a 89.5 = 30 da 90 a 100 = 30 e lode **Goethe-Zertifikat C1 =** 30 e lode **Goethe-Zertifikat C2 =** 30 e lode	
Tedesco (Test DaF)	**TEST DaF Institut** **media TDN = TDN3 =** 30 e lode **media TDN = TDN4 =** 30 e lode **media TDN = TDN5 =** 30 e lode	**TEST DaF Institut** **media TDN = TDN3 =** 30 **media TDN = TDN4 =** 30 e lode **media TDN = TDN5 =** 30 e lode	
Tedesco (Telc)	**Telc Deutsch B1 / +Beruf** meno di 180 punti = NON SUPERATO da 180 a 209,5 = ausreichend = 28 da 210 a 239,5 = befriedigend = 29 da 240 a 269,5 = gut = 30 da 270 a 300 = sehr gut = 30 e lode **Telc Deutsch B2 / +Beruf** 30 e lode **Telc Deutsch C1 / +Beruf, Telc Deutsch C1 Hochschule** 30 e lode	**Telc Deutsch B2 / +Beruf** meno di 180 punti = NON SUPERATO da 180 a 209,5 = ausreichend = 28 da 210 a 239,5 = befriedigend = 29 da 240 a 269,5 = gut = 30 da 270 a 300 = sehr gut = 30 e lode **Telc Deutsch C1 / +Beruf, Telc Deutsch C1 Hochschule** 30 e lode	
Tedesco (ÖSD)	**ÖSD Zertifikat Deutsch Österreich B1 / Zertifikat B1 Erwachsene/Jugendliche** meno di 60 punti = NON SUPERATO da 60 a 69 = ausreichend = 28 da 70 a 79 = befriedigend = 29 da 80 a 89 = gut = 30 da 90 a 100 = sehr gut = 30 e lode **ÖSD Zertifikat B2, ÖSD Zertifikat C1** 30 e lode	**ÖSD Zertifikat B2 Erwachsene/Jugendliche** meno di 60 punti = NON SUPERATO da 60 a 69.5 = 28 da 70 a 79.5 = 29 da 80 a 89.5 = 30 da 90 a 100 = 30 e lode **ÖSD Zertifikat C1** 30 e lode	

Livello minimo richiesto per il riconoscimento dell'esame	Quadro Europeo A1
Cinese – HSK (Chinese Proficiency Test)	**HSK 1** = 30 **HSK 2 – HSK 6** = 30 e lode

(1) Per le certificazioni Cambridge Assessment English precedenti il 2015, o che comunque non riportino il punteggio ottenuto secondo la "Cambridge English Scale", saranno utilizzate le seguenti equivalenze:

- **KET (A2)**
    - **Level A1 = 105**
    - **Pass/Grade C = 125**
    - **Pass with merit/Grade B = 135**
    - **Pass with distinction/Grade A = 145**
- **PET (B1)**
    - **Level A2 = 125**
    - **Pass/Grade C = 145**
    - **Pass with merit/Grade B = 155**
    - **Pass with distinction/Grade A = 165**
- **FCE (B2)**
    - **Level B1 = 145**
    - **Pass/Grade C = 165**
    - **Pass with merit/Grade B = 175**
    - **Pass with distinction/Grade A = 185**
- **CAE (C1)**
    - **Level B2 = 165**
    - **Pass/Grade C = 185**
    - **Pass with merit/Grade B = 195**
    - **Pass with distinction/Grade A = 205**
- **CPE (C2)**
    - **Level C1 = 185**
    - **Pass/Grade C = 205**
    - **Pass with merit/Grade B = 215**
    - **Pass with distinction/Grade A = 225**

(2) Tabella di equivalenza tra Cambridge IGCSE "First Language English" e livelli CEFR:
https://www.cambridgeinternational.org/Images/152745-cefr-levels-for-cambridge-igcse-first-language-english-0500-and-0522-.pdf

(3) Tabella di equivalenza tra Cambridge IGCSE "English as a Second Language" e livelli CEFR:
https://www.cambridgeinternational.org/images/152755-cefr-levels-for-cambridge-igcse-english-language-as-a-second-language-0510-and-0511-.pdf

Cambridge Assessment
International Education

CEFR Levels for Cambridge IGCSE® First Language English (A-G & 9-1 grading)

Common European Framework of Reference for Languages (CEFR)

Type of Language user	CEFR Level
Proficient user	C2 – Mastery
	C1 – Effective operational proficiency
Independent user	B2 – Vantage
	B1 – Threshold
Basic user	A2 – Waystage
	A1 – Breakthrough

CEFR levels of IGCSE E1L by skill

Syllabus 0500

	Reading	Writing		Integrated Listening & Speaking
Grade A	C2	C1	1	C1
Grade B	C1	C1	2	B2
Grade C	B2	B2	3	B1
Grade D	B2	B2	4	A2
Grade E	B2	B2	5	A2 or below
Grade F	B1	B1		
Grade G	B1	B1		

Syllabus 0990

	Reading	Writing		Integrated Listening & Speaking
7	C2	C1	Distinction	C1
6	C1	C1	Merit	B2
5	B2	B2	Pass	A2
4	B2	B2		
3	B2	B2		
2	B1	B1		
1	B1	B1		

Cambridge Assessment International Education

CEFR Levels for Cambridge IGCSE® English as a Second Language (A-G and 9-1 grading)

Common European Framework of Reference for Languages (CEFR)

Type of Language user	CEFR Level
Proficient user	C2 – Mastery
	C1 – Effective operational proficiency
Independent user	B2 – Vantage
	B1 – Threshold
Basic user	A2 – Waystage
	A1 – Breakthrough

CEFR levels of IGCSE E2L by skill

Syllabus 0511

	Reading	Writing	Listening	Speaking
Grade A	B2	B2	B2	C1
Grade B	B2	B2	B2	B2
Grade C	B2	B2	B2	B2
Grade D	B2	B2	B1	B1
Grade E	B1	B1	A2	B1
Grade F	A2	A2	A2	A2
Grade G	A2	A2	A2	A2

Syllabus 0991

	Reading	Writing	Listening	Speaking
7	B2	B2	B2	C1
6	B2	B2	B2	B2
5	B2	B2	B2	B2
4	B2	B2	B2	B2
3	B1	B1	A2	B1
2	A2	A2	A2	A2
1	A2	A2	A2	A2

Syllabus 0510 - with Speaking endorsed

	Reading	Writing	Listening		Speaking endorsed
Grade A	B2	B2	B2	1	C1
Grade B	B2	B2	B2	2	B2
Grade C	B2	B2	B2	3	B1
Grade D	B2	B2	B1	4	A2
Grade E	B1	B1	A2	5	A2
Grade F	A2	A2	A2		
Grade G	A2	A2	A2		

Syllabus 0993 - with Speaking endorsed

	Reading	Writing	Listening		Speaking endorsed
7	B2	B2	B2	Distinction	C1
6	B2	B2	B2	Merit	B2
5	B2	B2	B2	Pass	A2
4	B2	B2	B2		
3	B1	B1	A2		
2	A2	A2	A2		
1	A2	A2	A2		

Oxford Test of English Advanced

Test Specifications

April 2024

1   The Oxford Test of English suite

The Oxford Test of English is a suite of on-demand online English proficiency tests consisting of the *Oxford Test of English*, the *Oxford Test of English for Schools*, and the *Oxford Test of English Advanced*, developed by Oxford University Press (OUP) and certified by the University of Oxford. The tests are aligned to the Common European Framework of Reference for Languages (CEFR) (British Council, UKALTA, EALTA, & ALTE, 2022) and report across a range of CEFR levels (see Table 1). They are recognized globally as evidence of language proficiency.

Table 1. CEFR levels measured by tests in the Oxford Test of English suite

Test	CEFR level
Oxford Test of English Oxford Test of English for Schools	A2, B1, and B2
Oxford Test of English Advanced	B2 and C1

2   Introduction to Oxford Test of English Advanced

The *Oxford Test of English Advanced* measures at B2 and C1 levels on the CEFR for those who need to provide evidence of their English proficiency for educational and professional opportunities. These test specifications provide details of the design, content and skills tested in the *Oxford Test of English Advanced*, making explicit the links between the test design and the language skills identified as important for the purposes of the test.

3   Test description and rationale

The *Oxford Test of English Advanced* has been developed for users/learners of English to provide evidence of their English proficiency to a range of institutions, such as universities and employers, for gaining access to higher education and employment. The test content has therefore been designed to be suitable for students aged 16 and above. The test provides a valid and reliable measurement of English proficiency at B2 and C1 levels with tasks designed to assess the performances of the real-life tasks that plurilingual, pluricontextual language users/learners will meet in the educational and professional domains. Performance below level B2 is indicated as 'Below B2' in test results.

The *Oxford Test of English Advanced* focuses on English language users/learners' ability to both understand and communicate in English, as measured by four modules:

- Speaking
- Listening
- Reading
- Writing.

All modules are delivered entirely online and can be taken individually, or in any combination, on an on-demand basis.

The design of the test draws upon the CEFR Companion Volume (Council of Europe, 2020), reflecting the kinds of tasks and underlying skills that language users/learners encounter in academic and professional contexts. This includes an emphasis on mediation, reflected in the integrated skills tasks in the Listening, Speaking, and Writing modules, so that test score users can have confidence that test takers are able to

perform real-world tasks. The 'washback effect' (Alderson & Wall, 1993) on classroom teaching and learning was also taken into account when designing the test, with tasks selected which are likely to promote the development of skills appropriate for the academic and professional domains. OUP will be monitoring and reporting on the impact as data is collected from live test administration.

The content of the test is independent of any specific course of study. In preparation, test takers should familiarize themselves with the test, including marking criteria. Free preparation materials are available on the www.oxfordtestofenglish.com website. This includes:

- an online demo of the test
- downloadable practice tests
- a 15-hour online preparation course.

With respect to the number of guided study hours required to reach B2 or C1 levels, this will depend on a variety of factors such as age, motivation, teaching methodology, etc.

4   The test development process

The test was developed through an iterative design process (see Figure 1), aligned with the Association of Language Testers in Europe (ALTE) best practice, as set out in the Principles of Good Practice (ALTE, 2020) involving:

- initial test design based on the CEFR Companion Volume
- drafting of specifications
- production of sample materials
- reviews by internal assessment staff and external assessment consultants
- modification based on feedback and reviews
- trialling with students in teaching centres around the world
- test production
- pretesting
- analysis and review
- item banking
- CEFR alignment
- piloting the test at approved test centres around the world.

Figure 1. The test development process
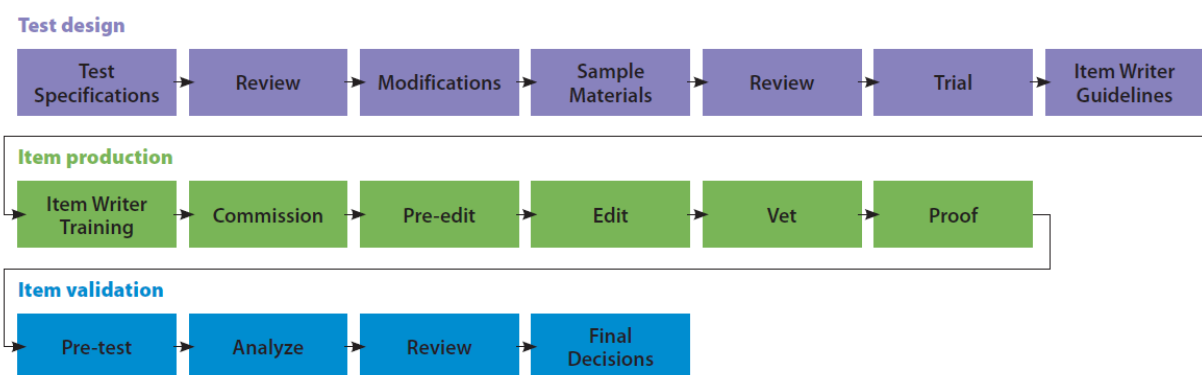
The first phase of test development involves producing comprehensive test specifications, culminating in this document. The initial specifications detailed the desired test format, proposed content for each of the modules and each of the tasks contained within them. Well-crafted specifications communicate the test designers' vision, underpinning the consistency of measurement (i.e. reliability) across modules, enhancing the quality of the test across administrations and help to ensure that decisions made based on test scores will be fair and valid.

In creating the specifications for the *Oxford Test of English Advanced*, OUP worked closely with institutions, teachers and learners to ensure that the test met their needs, while making certain that the test was also aligned to OUP's approach to language teaching, learning and assessment.

The specifications for the *Oxford Test of English Advanced* were derived from:

- level and domain descriptions in the CEFR Companion Volume (Council of Europe, 2001): each task in the test is related to one or more CEFR Companion Volume 'Can Do' descriptors, including new descriptors for mediation where relevant. Mediation descriptors influenced the design process and led to the introduction of several integrated-skills tasks within the test
- communicative teaching practice
- course outlines and content from OUP teaching materials.

Mediation plays a central role in the test design in recognition of its crucial role in communication. The CEFR divides communication into four modes: reception, production, interaction, and mediation. Reception refers to activities such as reading an article or listening to a lecture, where someone receives and processes input. Production refers to activities, such as giving a talk or writing an essay, where someone speaks or writes at some length. Interaction, such as conversation, involves two parties co-constructing discourse. Mediation, which reflects much of our communication, is a combination of two or all of reception, production and interaction. Examples in the *Oxford Test of English Advanced* include Listening Part 2, which involves reading, listening and note taking; Speaking Part 3 (Summary) which includes listening, note-taking and speaking; and Writing Part 2 (Summary) which involves reading, note-taking and writing.

The test is designed to cover as wide a range of domains as possible within the confines of a two-hour administration. Independent language-testing professionals were invited to comment on the draft specifications to help ensure appropriate coverage of domains and levels. These draft specifications were reviewed by an internal OUP panel and revised ahead of the production of sample materials. The specifications were then reviewed a second time, along with these sample materials, and further modifications were made. The specifications then formed the basis of the item writer guidelines (IWGs). These documents are the procedural outlines for item writers on how to write items for each of the task types of the test.

Item-writer Team Leaders were commissioned to draft item writer guidelines for each module, based on the specifications and sample materials. These guidelines help our item writers to produce comparable, high-quality tasks to ensure consistency across different instances of the test and to ensure that tasks continue to reflect the intentions of the designers. The item writer guidelines are used in tandem with OUP's Diversity, Equity and Inclusion (DEI) guidelines which promote inclusive content. A team of item writers was trained to write an initial set of test materials. These fed into small-scale trialling in which groups of students were asked to take these tasks and provide feedback on the experience. Another

round of minor revisions was then made based on the comments from other item writers and from trial students. Further sets of materials were then commissioned. These were pretested more extensively on representative samples of students in a range of countries worldwide.

5   Oxford Test of English Advanced overview

Module	Part	No. tasks	No. items per task	Structure	Timing
Speaking	1	1	6	Interview: six questions on various everyday topics	Approximately 15 minutes
	2	1	1	Voice message: respond to a situation requiring diplomacy	
	3	1	1	Summary: summarize and synthesize the main ideas expressed by two different expert speakers	
	4	1	1	Debate: put a case for or against a proposition	
	5	1	4	Follow-up questions: respond to four questions related to the debate	
Listening	1	5	1	3-option multiple-choice questions on independent monologues/dialogues	Approximately 40 minutes
	2	1	5 or 6	Longer monologue with a note-completion task	
	3	1	5 or 6	Longer dialogue matching speakers to opinions	
	4	5	1	3-option multiple-choice questions on independent monologues/dialogues	
Reading	1	6	1	3-option multiple-choice questions on a variety of text types	Approximately 35 minutes
	2	1	6 or 7	Multiple-matching task	
	3	1	6	Gapped text task with extracted sentences	
	4	1	4 or 5	Multiple-choice questions on a longer text	
Writing	1	1	1	Essay	50 minutes
	2	1	1	Summary	

6   Module overviews

6.1   Speaking module overview

In the Speaking module, test takers wear a headset and speak into a microphone to answer questions delivered by computer. Input is either audio-only (i.e. the text of the task is heard, but not shown on screen) or audio-written (i.e. the text of the task is heard and shown on screen). An onscreen clock shows how much time is available to answer each question. Preparation time is given for the Voice message in Part 2, the Summary in Part 3 and the Debate in Part 4 before the test taker has to begin speaking.

Part	No. tasks	No. Items per task	Structure	Testing focus
1	1	6	**Interview** Answering six spoken single-sentence questions Questions 1 and 2 always the same and are given to all test takers Questions 3–6 **Audience:** the interviewer / assessor **Preparation time:** none **Response time:** 30 seconds per question	• expressing opinions and attitudes • speculating about the past and future • hypothesizing • giving elaborate descriptions • describing the personal significance of events and experiences • advising
2	1	1	**Voice message** Responding diplomatically to a difficult or sensitive situation in a voice message **Audience:** a colleague in a professional or academic setting **Preparation time:** 10 seconds **Response time:** 40 seconds	• expressing disagreement diplomatically • negotiating complex or sensitive transactions
3	1	1	**Summary** Summarizing and synthesizing the main ideas expressed in two audio input texts on a complex subject **Audience:** colleagues in an academic setting **Preparation time:** 40 seconds **Response time:** 50 seconds	• identifying relevant information • distinguishing main points and supporting details • synthesizing information from multiple sources • paraphrasing and circumlocution
4	1	1	**Debate** Putting a case for or against a proposition in a debate **Audience:** colleagues in an academic setting **Preparation time:** 45 seconds **Response time:** 2 minutes	• arguing a case on a complex issue • structuring an argument • formulating points precisely • employing emphasis effectively
5	1	4	**Follow-up questions** Responding to questions expanding on the topics raised in the debate **Audience:** the interviewer / assessor **Preparation time:** none **Response time:** 40 seconds per question	• speculating and hypothesizing • making critical remarks

6.2 Listening module overview

The timing of all parts of the Listening module is predetermined. In each part, test takers hear each recording twice, with the option of moving to the next task after the first audio play completes, and are given a set time to check their answers before the test automatically progresses to the next recording.

Part	No. tasks	No. Items per task	Structure	Testing focus
1	5	1	Five short monologues / dialogues each with one 3-option multiple-choice	• identifying gist • identifying function/reason/purpose

			question (text options or picture options*)	• understanding attitude/feeling/opinion • identifying type • identifying speaker relationship • identifying topic • understanding implied meaning • understanding rhetorical purpose (exemplification, comparison, exaggeration) • identifying specific information (identifying finer points of detail)
2	1	5 or 6	Note completion A longer informational / descriptive monologue with note-completion task (open cloze OR 3-option MCQ*)	• identifying specific information
3	1	5 or 6	Matching opinions to the correct person / people A longer formal or semi-formal dialogue with five or six 3-option multiple-choice questions	• identifying stated opinion and implied meaning
4	5	1	Five short monologues / dialogues each with one 3-option multiple-choice question (text options)	• as for Part 1

*Task type and number of items the test taker receives will depend on the ability estimate generated by their response to the previous question.

6.3   Reading module overview

Texts used in Parts 1, 3 and 4 in the Reading module are based on authentic material intended to be of relevance or interest to a general readership. Texts for Part 2 are inspired by authentic reading sources and written by item writers. Authentic sources are acknowledged at the end of the test. Texts may be formal, neutral or informal in register. The time allowed for completion of each task in the Reading module is predetermined. If the test taker does not complete the task within the allotted time, the system will automatically progress to the next task.

Part	No. tasks	No. Items per task	Structure	Testing focus
1	6	1	Six short texts, each with one multiple-choice question	Careful reading at local and global levels: identifying main message, purpose, detail and implied meaning
2	1	6 or 7	Six profiles of people to match with four texts OR Seven items to match with three texts	Expeditious search reading at local and global levels: identify specific information, opinion and attitude and implied meaning
3	1	6	Gapped text, each with six extracted sentences plus one distractor	Careful global reading: identifying text structure and organizational features of a text
4	1	4 or 5	Four or five multiple-choice questions on a longer text	Propositional and rhetorical purpose items: for example, identifying attitude/opinion, purpose, understand implication and exemplification

6.4 Writing module overview

In both parts, test takers type their responses. The tasks specify a target audience and a minimum and maximum word count. There is an automatic word count facility. Marks are capped if responses are under length (Part 1 only) or overlength (Part 2 only).

There is a clock so that test takers always know how much time they have remaining for each part.

Part	No. tasks	No. Items per task	Structure	Testing focus
1	1	1	**Essay** Writing an essay in response to an essay question plus topic prompts which must be included in the response **Audience:** academic tutor **Response length:** 220–280 words **Response time:** 30 minutes	• structuring an argument with suitable introduction and conclusion • giving reasons in support of or against a particular point • evaluating different ideas or solutions to a problem • highlighting salient issues
2	1	1	**Summary** Reading two texts then writing a single-paragraph summary combining information from the texts **Audience:** peers in an academic setting **Response length:** 80–100 words **Response time:** 20 minutes	• identifying relevant information • distinguishing main points and supporting details • synthesizing information from multiple sources • paraphrasing and circumlocution

7 Test delivery

Unlike traditional 'linear tests' (such as paper-based tests or linear computerized tests,) the *Oxford Test of English Advanced* does not have fixed test versions in which all test takers encounter the same set of questions. Instead, it operates using an item bank and a series of selection rules. An item bank is a large collection of test questions (also known as 'items') from which individual questions are presented to individual test takers. The large number of test questions helps to ensure that different test takers using the test at the same time receive different sets of questions. To ensure that each test taker has the same test experience, the *Oxford Test of English Advanced* delivers a standard test format to each test taker. That is, all test takers receive the same task types and the same number of items. The final ability estimate is derived once the complete set of test items in a module has been delivered. Ability estimates are then converted to a standardized score, and this is also reported in terms of a CEFR level.

As test takers do not receive the same set of items, test security is improved, allowing the *Oxford Test of English Advanced* to be used on an on-demand basis, rather than limiting delivery to scheduled sessions. And, as the test is delivered wholly online, no materials need to be transported to test centres and stored on site, which also increases security. The item bank is refreshed on a regular basis to ensure that items do not become over-exposed. Finnerty (2015) gives further details about the advantages and workings of computer-adaptive testing.

7.1 Computer-based testing

The *Oxford Test of English Advanced* is a modular online computer-based test. The computer-based approach has a number of advantages over more traditional testing, including a shorter test length, increased precision of measurement, a wider range of task types, greater support for the test taker, and

improved accessibility features. For example, the *Oxford Test of English Advanced* includes timed reading tasks and so can differentiate between speeded (or 'expeditious') reading and careful reading, which requires more time. Other examples include the integrated tasks which focus on mediation, such as the listening into speaking summary task, and the Voice message task, which focuses on pragmatic meaning. Two approaches are taken in computerized testing: computer-adaptive testing, and randomized testing.

7.1.1   Computer-adaptive modules

The Listening and Reading modules are computer-adaptive, an algorithm-based scoring model which provides more precise measurement, and much shorter tests, than equivalent traditional tests. The algorithm, which determines the test taker's ability estimate, works in combination with test selection rules which determine which items are presented to each test taker, for example, 'choose five Part 1 Listening tasks'. A randomness element is also factored into the selection of tasks, so that each test taker receives their own individualized version of the test.

An item bank was created consisting of item sets at different Rasch difficulty values (Bond, Yan & Heene, 2020), at B2 and C1 level. The difficulty values are determined through statistical analysis during the pretesting stage. Pretesting is a type of 'try out' (ALTE, 2011) of items, where linear tests containing uncalibrated items are taken, along with an 'anchor test' of pre-calibrated items, under test conditions on a representative cohort. For the *Oxford Test of English Advanced*, a pretest consists of items for a single skill (either Listening or Reading) and typically follows the format of the live test. Pretests are taken at approved research partner institutions (typically language schools). Sampling guidelines set out the pretesting parameters (for example, a minimum of 150 test takers across a minimum of three language groups). Pretest cohorts are drawn from a range of populations representative of the test taking population to ensure the appropriateness of the those likely to take the test. Test taker response data is then analyzed to determine the Rasch difficulty of each item. Qualitative questionnaire data on the test taker's perception of the test is also analyzed. Items which fail to meet acceptance criteria are rejected or amended and re-pretested.

During the test, the algorithm draws on the item bank and presents the first item to the test taker. The test taker answers the items in the item set and the computer automatically marks them and produces an ability estimate based on the number of items in the set that were answered correctly. When a test taker answers sufficient number of items correctly, the algorithm revises the test taker's ability estimate upwards and the next item selected will therefore be more difficult; when a test taker answers insufficient items correctly, the algorithm revises the test taker's ability estimate downwards and the next item selected will be easier. The test continues with the algorithm selecting each subsequent item depending on the test taker response to the preceding items, and the estimate of the test taker's ability is refined using additional information from each item or item set and the statistical error associated with the estimate is reduced.

7.1.2   Computer-randomized modules

The Speaking and Writing modules are randomized linear tests. The Speaking and Writing item banks contain tasks written to be accessible to, and elicit responses at, CEFR B2–C1. During the test, tasks are randomly selected for each part of the test from the item banks, meaning that different test takers within a test session will receive different tasks to other test takers. Test taker responses are marked by

trained human assessors against analytic criteria. Separate criteria are used to mark Part 1 (Essay) and Part 2 (Summary) (See Appendices 6 and 7 respectively).

7.2   Test administration

The *Oxford Test of English Advanced* can only be administered by Approved Test Centres (ATCs). Test centres have to provide evidence that they meet technical requirements and have the appropriate facilities and suitable staff to administer the test securely. ATCs are subject to ongoing quality control checks and audits. Requirements for test centres are detailed in the *Oxford Test of English Test Centre Handbook* (see https://elt.oup.com/feature/global/oxford-test-of-english/documents?cc=gb&selLanguage=en).

Once approved, a test centre can purchase test licenses to run the test. The test centre then selects the date or dates on which they wish to run the test and allocates licenses to that session. The *Oxford Test of English Advanced* can be taken on any date, though OUP usually requires seven days' notice of a test session – this ensures that sufficient assessors are allocated for the marking of Speaking and Writing modules. The *Oxford Test of English Advanced* is usually taken as a complete test (all four modules are administered in the course of a session), but test centres may choose to run sessions for single modules or any combination of modules. It is also possible for test takers to choose to resit individual modules, rather than resitting the whole test.

7.3   Reasonable adjustments

Oxford University Press is committed to providing reasonable adjustments to make the *Oxford Test of English Advanced* accessible to users/learners with disabilities wherever possible. A range of support options related to visual difficulties, hearing difficulties, learning difficulties or access needs are available. To request help, the test centre manager must complete a Reasonable Adjustments Request Form, available at www.oxfordtestofenglish.com, on behalf of the test taker.

The following reasonable adjustments must be requested six weeks in advance supported with medical evidence:

- extra time
- rest breaks between modules
- individual test session
- physical testing environment
- access to the test room
- assistive role: Scribe
- assistive role: Practical assistant
- assistive role: Reader
- assistive role: Prompter.

The following reasonable adjustments do not need to be requested in advance. They can be made by the test taker during the test.

- display option: enlarged text
- display option: low contrast (pastel)
- display option: high contrast
- reading support device

- keyboard controls
- hand-held text magnifier.

8 Marking and scoring

8.1 Listening and Reading modules

Responses for Listening and Reading are marked by computer and the ability of the test taker is estimated according to the responses given in relation to the difficulty of the questions presented. The *Oxford Test of English Advanced* employs the Weighted Maximum Likelihood Estimation (Warm, 1989) in its test algorithm. The equation in this formula uses the test taker's responses to items of different Rasch difficulties to estimate their ability at each decision point, i.e. at the end of each item or set of items. The final Rasch ability estimate is converted to a standardized score on the Oxford Test of English Scale.

8.2 Speaking and Writing modules

Speaking and Writing tasks are selected at random from the item bank, according to a pre-defined number and order of tasks, and the responses are returned online and sent electronically to trained assessors who mark them according to analytic criteria (also known as 'rubrics') derived from the CEFR level descriptors. Analytic criteria are used as they ensure that assessors focus on a range of marking elements rather than focus too heavily on one area of the test taker's performance, as can be the case in holistic criteria.

Speaking criteria consist of Task fulfilment, Pronunciation and fluency, Grammar, and Lexis. The table below summarizes the main elements of the marking criteria. See Appendix 5 for detailed Speaking marking criteria.

Table 2. Elements of the Speaking marking criteria

Task fulfilment	Pronunciation and fluency	Grammar	Lexis

• Task requirements
• Impact on listener
• Register
• Synthesis (Part 3 only)
	• Stress, rhythm, and intonation
	• Flow
	• Organization:
	  o Coherence
	  o Cohesion
		• Control and accuracy
		• Range
			• Range and accuracy
			• Collocations, idioms, and colloquialisms

Writing has separate criteria for Part 1 (Essay) and Part 2 (Summary) to reflect the different skills required for these tasks. Both criteria consist of Task fulfilment, Organization, Grammar, and Lexis, but have different foci for the sub-criteria. The tables below summarize the main elements of the marking criteria. See Appendices 6 and 7 for detailed Writing marking criteria.

Table 3. Elements of the Writing marking criteria – Part 1 (Essay)

Task fulfilment	Organization	Grammar	Lexis
• Essay development • Register • Impact on reader	• Organization of ideas • Coherence • Cohesion	• Range and control of structures • Accuracy	• Range • Appropriacy • Accuracy

Table 4. Elements of the Writing marking criteria – Part 2 (Summary)

Task fulfilment	Organization	Grammar	Lexis
• Synthesis of information • Clarity of communication • Register	• Organization of ideas • Coherence • Cohesion	• Independence from input structures • Control and conciseness • Accuracy	• Independence from input lexis • Control and conciseness • Accuracy

The criteria are on an eight-point scale (0–7) covering B1 to C2. There are detailed descriptors for B1, B2, C1 and C2 which represent the criteria required to demonstrate performance at the relevant CEFR level. These descriptors apply to performance at both the *criterion* level and the *plus* level for the CEFR level, e.g. the B2 descriptors (which are presented under B2.1) relate to performance at both B2.1 and B2.2. To decide whether a performance is at the *criterion* level or *plus* level, the assessor decides if all the positive descriptors have been met *minimally* or *comfortably*, and if any negative descriptors have been met *minimally* or *not at all*. If all positive descriptors have been met *comfortably* and any negative descriptors have been met *minimally* or *not at all*, a *plus* level is awarded. Assessors are provided with exemplar responses that they can refer to for examples of *criterion* and *plus* level performance at each CEFR level. Whilst the marking criteria cover B1 to C2, results are only reported at B2 and C1 levels. The B1 and C2 descriptors are used as some C1 users/learners may demonstrate aspects of these descriptors in their responses, allowing a greater range of marks to be awarded. The test does not award B1 or C2 grades as the tasks presented have not been designed for this purpose.

The length and relevance of the test taker response is taken into account when awarding marks. In the Speaking module, different penalties are applied depending on the extent and relevance of the response. In the Writing module, caps are in place depending on the extent and appropriacy of the response. See the marking criteria for further details. Test taker responses are anonymized and split into two 'scripts', as shown in the table below, and each script for a module is marked separately. The marks of the two scripts are combined and converted to a standardized score and CEFR level for each module.

Table 5. Scripts in the Speaking and Writing modules

	Script 1	Script 2
Speaking module	Part 1 (Interview) Part 2 (Voice message) Part 3 (Summary)	Part 4 (Debate) Part 5 (Follow-up questions)
Writing module	Part 1 (Essay)	Part 2 (Summary)

8.3   Assessor marking quality assurance

All Speaking and Writing assessors have significant English language-teaching experience and recognized English language-teaching qualifications. Assessors follow a standardized training and certification process before being allowed to participate in marking. Their marking is then monitored to ensure consistency. Automated quality assurance monitoring is carried out using Speaking and Writing responses which have been marked previously by a number of experienced assessors and so have agreed benchmark ratings. These 'seeded' responses are interspersed with test taker responses to check that the assessors continue to be accurate in their marking to within set tolerances. All responses are anonymous, so assessors are unaware whether the responses they are marking are test taker responses

or seeded responses. Assessors whose marking falls outside of agreed tolerances are removed from the marking process and asked to complete a re-standardization process, after which they can resume marking. Assessors who do not successfully complete re-standardization are permanently withdrawn from marking.

9 Results reporting

Performance on the *Oxford Test of English Advanced* is reported in terms of standardized scores on a scale ranging from 0 to 170 and related CEFR level. Standardized scores are independent of test sessions and give a standard reference point for students taking the test on different occasions. Results are also displayed as a bar chart, showing how performance on the test relates to the relevant CEFR levels. The table below shows the relationship between the *Oxford Test of English Advanced* scale and the CEFR.

Table 6. CEFR bands and standardized score ranges for the Oxford Test of English Advanced

CEFR band	Oxford Test of English Advanced standardized score range
C1	141–170
B2	111–140
Below B2	0–110

For the objectively marked Reading and Listening modules, the final ability estimates obtained through the test algorithm are converted to standardized scores and these are used in determining the CEFR levels. For the assessor-marked Speaking and Writing modules, marks for each task in the module are combined and converted into standardized scores.

Test takers receive a standardized score and CEFR level on a Module Report Card for each module taken. If a test taker completes all four modules, they also receive an overall score and CEFR level on an Oxford Test of English Certificate. The overall score is calculated as an average of the scores obtained in each of the four modules. See Figure 2.

Test takers receiving a standardized score under the B2-level threshold do not receive a CEFR mark, but an indicative 'Below B2'. The rationale for this is that the test taker has received tasks designed for CEFR levels B2 and C1, so we cannot be confident how they would have performed on tasks designed for A2 and B1 test takers. Below B2-level performance means that a test taker is not at the level the test was designed to measure and that no precise statement of CEFR level can be made.

Results for the Reading and Listening modules are available on the same day that the test was taken. Results for Speaking and Writing are ordinarily available in five days.

Figure 2. Sample Oxford Test of English Advanced Module Report Card for the Speaking module, and sample Oxford Test of English Advanced Certificate

10 Results interpretation and use

The *Oxford Test of English Advanced* has been designed for young adult and adult test takers aged sixteen or above as evidence of English proficiency for the purposes of gaining entry to education and employment opportunities. Results can be verified via the Verification of Results page (https://go.oup.com/oxfordtestofenglish/verify) and are valid for life. However, universities, employers, or other institutions may wish to specify that results have been gained within a time limit, for example, in the last two years. Test takers should ensure that they are aware of the requirements of the institutions they are applying to. OUP reserves the right to amend or withdraw results following a results review, appeal or other results-related investigation.

11 Results reviews and appeals

However effective a testing program may be, test takers or other stakeholders may wish to challenge or appeal their result and transparent procedures must be open to them. There is a two-stage process for challenging a result on the Oxford Test of English: results review and appeal.

For a results review, the test results for one or more modules are checked or re-marked. For Speaking and Writing, a results review involves a re-mark of the responses. This is done by inviting senior assessors to re-mark the module in question. If the re-mark results in a score that improves the module or overall CEFR level, the results enquiry is upheld and the test taker receives a replacement result.

For Listening and Reading, the results review will involve a results check. As Listening and Reading are both marked by computer, there is no scope for re-marking as the re-mark result would be identical to the original result. However, a check is made by OUP on the tasks presented to the test taker to ensure that they received tasks at the appropriate level and that their ability estimate was correctly calculated. If an error is identified with the result, a decision will be made as to whether a revised result can be issued or whether the test taker should be given the opportunity to resit the module.

A test taker can also request an appeal via their test centre. An appeal differs from a results review in that an appeals panel, which is entirely independent of OUP, undertakes the investigation of the test taker's responses and marks to ensure that all appropriate steps have been taken in reviewing the result. The Oxford University Department for Continuing Education (OUDCE) acts as the independent appeals body for the Oxford Test of English. An appeal must be preceded by a results review.

An administrative fee is charged for all results reviews and appeals, but the fee is refunded if the review results in a change of CEFR level for either a module or the whole test, or if the appeal is upheld. All results reviews and appeals are processed on behalf of the test taker by the test centre at which the test was administered.

12 Test monitoring, impact and review

The development and administration steps outlined above have been designed to ensure that every administration of the *Oxford Test of English Advanced* provides reliable results that serve as a valid basis for decision-making.

To ensure that the *Oxford Test of English Advanced* continues to fulfil its stated purpose, and to seek opportunities for further improvements in quality, OUP monitors test administrations and carries out analyses of the performance of test materials, test takers, test centres and assessors at regular intervals. This includes differential item functioning to investigate potential bias based on sex or first language.

As part of the University of Oxford's recognition, endorsement and certification, OUDCE reviews the performance of the test on a quarterly basis to ensure that it meets quality standards.

13 Test security

OUP takes test security seriously and a number of measures, both overt and covert, are taken to protect the integrity of the test. For example, the test can only be taken at Approved Test Centres that undergo a rigorous onboarding process and are regularly monitored and inspected. Test sessions are invigilated by trained invigilators, and the test itself is taken in a lockdown browser which prevents test takers from accessing any applications, such as web browsers. Regular statistical analysis is carried out on test taker response data and test centre activities to identify potential irregularities.

14   Appendix 1: Speaking module details

14.1   Speaking module construct and domain definitions

Test purpose	The *Oxford Test of English Advanced* Speaking module assesses test takers' oral English proficiency for three sub-domains: academic, professional and social purposes. Within these sub-domains, the test assesses the ability to:
	• express personal opinions and provide information when responding to questions about personal experiences and preferences on familiar topics
	• negotiate and disagree diplomatically in more formal and professional settings
	• identify main ideas, synthesize and summarize them from listening materials from an academic topic from typical introductory lectures
	• sustain and structure an argument for a debate which employs emphasis on key messages
	• respond to questions similar to those which could be encountered in a typical academic-style seminar.

Examples of domain activities	**Academic domain:**
	• **Participating in an academic debate**: responding to prompts related to a classroom discussion, expressing opinions on academic topics, and disagreeing diplomatically with classmates, speaking at length on a single topic, arguing a point of view
	• **Presenting research findings:** providing a concise response to prompts about presenting research findings, conveying complex information within the time limit
	• **Professional domain:** job interview simulation, client interactions and inquiries
	• **Social domain:** making plans, expressing opinions, resolving conflicts, expressing feelings and opinions in a conflict situation, disagreeing diplomatically, maintaining positive social relationships.

Communication goals by sub-domain		Academic	Professional	Social
	Giving information	☑	☑	☑
	Responding to questions	☑	☑	☑
	Responding and expressing opinions appropriately	☑	☑	☑
	Employing emphasis effectively	☑	☑	☑
	Arguing a case on a complex issue	☑	☑	☐
	Structuring an argument	☑	☑	☐
	Formulating points precisely	☑	☑	☐
	Reinforcing key messages	☑	☑	☐
	Synthesizing ideas	☑	☑	☐
	Summarizing an argument	☑	☑	☐
	Disagreeing diplomatically	☑	☑	☐
	Negotiating complex or sensitive transactions	☑	☑	☐

Targeted cognitive functions	**Planning and goal orientation:** the test taker's use of strategies to maximize communicative success with available resources, and to compensate for limitations on their speaking performance.
	Formulation: using linguistic resources such as vocabulary (lexicalization) and grammar (grammatical encoding) to express meaning.
	Articulation: pronunciation of words and phrases intelligibly and ability to use intonation, stress and rhythm to express meaning.
	Discourse organization: conceptualizing, connecting and developing ideas effectively to convey message comprehensibly and accurately.

	Monitoring: the test taker's awareness of effective ways to prepare themselves for spoken communication in the target language and context; awareness of breakdown and attempts at repairs.

14.2 Speaking Part 1 (Interview) design and sample tasks

Module	Speaking
Part	1
Task label	Interview
CEFR	B2–C1

14.2.1 Speaking Part 1 task features

Please note that Speaking tasks can be:

- audio-only: the test taker hears the task, but does not see the task (only the generic rubric is presented visually on screen)
- audio-written: the test taker both sees and hears the task.

Skill focus	• Responding to questions • Expressing opinions appropriately
Description	Test takers hear and respond to six audio-only questions. Questions 1 and 2 target biodata, and responses are not assessed. Biodata questions are not commissioned. Questions 3–6 are thematically unrelated and are of a personal nature and are designed to elicit a range of extended responses (see below for examples).
Number of items per task	Four thematically unrelated questions.
Time allocated for task	Response time: 30 seconds per question (2 minutes total).
Instructions for candidates (task rubric)	You are going to answer six questions. The clock shows how much time you have to speak. Start speaking when you hear the tone. Try to speak for the full amount of time.
Question format and presentation	Test takers are presented with an audio-written standard rubric. The task questions are audio-only, supported by the (written-only) number of the question appearing on the screen (e.g. 'Question 1') for the duration of the time allotted for the question being asked, plus the time remaining for the question shown in an onscreen clock. Question numbers do not form part of the audio.
Targeted CEFR level(s)	B2– C1
Targeted cognitive functions	☑ Monitoring
	☑ Articulation
	☑ Buffer
	☑ Formulation
	☑ Planning
	☑ Conceptualization

14.2.2 Speaking Part 1 features of the input

Description	- The questions should be ordered from least cognitively demanding to most cognitively demanding. - The first question should begin: 'Thinking about (topic)… '. For example, "Thinking about your friends, tell me about something you have done with your friends recently." - The final question should begin: 'Finally … '. For example, "Finally, how do you feel about the future, worried, or excited?" - Each question should have a different focus.
Word count	For each set of four questions the total word count is 60 words (excluding the task rubric).
Grammatical and/or phonological features	- Mixture of direct and indirect questions (single or multi-clause) - Questions avoid all but the most basic contractions. Questions may be fronted with buffers (e.g. 'Can you tell me what you have done with your friends recently.', **NOT** 'What have you done with your friends recently?').
Content (specificity, neutrality, concreteness)	Concrete/specific
Lexical range (refer to Oxford 5000 for CEFR levels of individual words)	Primarily A1–B2 lexis. Questions are written to be accessible to B2-level test takers but allow C1 test takers to demonstrate their proficiency.
Genre/domain (e.g. personal/professional/education/other)	Personal

14.2.3 Speaking Part 1 features of the expected response

Description	The task is designed to elicit responses around the test taker's personal preferences and everyday experiences.
Presentation (written/recorded/illustrations/options)	Verbal; audio-recorded. After each question is heard, audio recording commences immediately. A timer is visible, counting down from 30 seconds. Recording automatically ceases once timer reaches zero seconds.
Discourse functions	☑ Giving information
	☑ Describing
	☑ Comparing
	☑ Narrating
	☑ Hypothesizing
	☑ Expressing opinions and attitudes
	☑ Describing dreams, hopes and ambitions
	☑ Speculating about the past and future
Scoring model	Part 1 test taker responses are marked alongside Parts 2 and 3 as a whole, referred to as 'Script 1'. Test takers receive four scores, one mark per criterion: Task fulfilment, Pronunciation and fluency, Grammar, and Lexis. These are added together with marks from Script 2 (Parts 4 and 5). The total marks are then translated into a standardized score and CEFR level.
Supplementary information/criteria	Questions should be straightforward but elicit a range of possible language structures.

14.2.4 Speaking Part 1 sample task

This task is an extract from an *Oxford Test of English Advanced* practice test. To see how the task displays on screen in the test, see the next section.

You are going to answer six questions.
The clock shows how much time you have to speak.
Start speaking when you hear the tone.
Try to speak for the full amount of time.

1. What's your name?
2. Which country do you come from?
3. Thinking about education, what would you like to study in the future?
4. What's the best time of year to visit your country?
5. How have your interests changed as you have got older?
6. Finally, how important are books in your life?

14.2.5 Speaking Part 1 screen display

Individual questions are heard but not displayed on the screen.

14.3 Speaking Part 2 (Voice message) design and sample tasks

Module	Speaking
Part	2
Task label	Voice message
CEFR	B2–C1

14.3.1 Speaking Part 2 task features

Skill focus	• Negotiating complex or sensitive transactions • Giving information • Expressing opinions and feelings • Disagreeing diplomatically
Description	Test takers are presented with an audio-written situation. Using three bulleted prompts, the test taker produces a 40-second voice message response.
Number of items per task	One
Time allocated for task	Preparation time: 10 seconds Response time: 40 seconds
Instructions for candidates (task rubric)	You are going to leave a voice message. First read and listen to the task, then decide what you want to say. You need to be diplomatic in your response. The clock shows how much time you have to leave your voice message. Start speaking when you hear the tone. [Task details (see 'Question format and presentation' and 'Description' below)] You now have some time to think about what you want to say.
Question format and presentation	Test takers are presented with an **audio-written** standard rubric, comprising the instructions (see above) and the task details: You [study at college / work for a company]. [task-specific text]. Leave a voice message for [audience] and: • [task-specific prompt 1] • [task-specific prompt 2] • [task-specific prompt 3].
Targeted CEFR level(s)	B2– C1
Targeted cognitive functions	☑ Monitoring
	☑ Articulation
	☑ Buffer
	☑ Formulation
	☑ Planning
	☑ Conceptualization

14.3.2 Speaking Part 2 features of the input

Description	There are five elements to the Voice message tasks: **Element 1: Domain.** The domain (academic or professional) is set in the first sentence. This must be worded using one of the following: Academic domain: *You study at college.* Professional domain: *You work for a company.* **Element 2: Relationship to the recipient.** The voice message recipient must be related to the domain established in Element 1. The relationship with the recipient must be established (e.g. 'your colleague', 'your manager', 'your tutor', 'a student at the college', etc.); the relationship signals register (i.e. level of formality) required in the test taker's response. **Element 3: Situation.** A situation for the test taker to respond to diplomatically. Examples include: a situation that the test taker wants the recipient to change (Example 1); a request that the test taker wants to refuse (Example 2), etc. The situation can be formal or informal. **Element 4: Audience.** An audience must be specified for the test taker to leave a voice message for. The audience will be a person, not an institution or organization. The audience will typically also be mentioned in the situation. **Element 5: Prompts.** There will be three prompts arranged as bullet points.
Word count	70 words (excluding the question rubric).
Grammatical and/or phonological features	The task format is fixed, requiring the grammatical structures identified in 'question format and presentation' and the 'description' above.
Content (specificity, neutrality, concreteness)	Concrete/specific
Lexical range (refer to Oxford 5000 for CEFR levels of individual words)	Primarily A1–B2 lexis. Questions are written to be accessible to B2 level test takers but allow C1 test takers to demonstrate their proficiency.
Genre/domain (e.g. personal/professional/education/other)	Academic/professional

14.3.3 Speaking Part 2 features of the expected response

Description	The task is designed to elicit an extended response requiring diplomacy. The response will be approximately 80–90 words in length. Test takers should aim to use their own words where possible, avoiding using too much language from the task input.
Presentation (written/recorded/illustrations/options)	Verbal; audio-recorded. After question is heard, audio recording commences immediately. A timer is visible, counting down from 40 seconds. Recording automatically ceases once timer reaches zero seconds.
Discourse functions	Organizing and sustaining extended discourse
	Responding and expressing oneself appropriately in different professional or academic situations
	Sustaining relationships using appropriate politeness conventions

	Sustaining transactional discourse in professional or academic situations
Scoring model	Part 2 is marked holistically alongside Parts 1 and 3 (as part of Script 1). Test takers receive four scores. Scores for 'Task fulfilment', 'Pronunciation and fluency', 'Grammar' and 'Lexis'. These are added together with scores from Script 2 (Parts 4 and 5). The total score is then translated into a whole CEFR level.

14.3.4 Speaking Part 2 sample task

This task is an extract from an *Oxford Test of English Advanced* practice test. To see how the task displays on screen in the test, see the next section.

Voice message

You are going to leave a voice message. First read and listen to the task, then decide what you want to say. You need to be diplomatic in your response.
The clock shows how much time you have to leave your voice message. Start speaking when you hear the tone.

You work for a company. Your manager, Paul Chapman, has asked you to give an important presentation to some clients. You are unable to give the presentation. Leave a voice message for your manager and:

- say why the presentation is important
- explain why you are unable to give the presentation
- suggest who should give the presentation, and why.

You now have some time to think about what you want to say.

14.3.5 Speaking Part 2 screen display

14.4 Speaking Part 3 (Summary) design and sample tasks

Module	Speaking
Part	3
Task label	Summary
CEFR	B2–C1

14.4.1 Speaking Part 3 task features

Skill focus	• Integrated skills; listening into speaking • Giving information • Summarizing an argument
Description	Summarizing and synthesizing the main ideas expressed in two audio input texts on a complex subject.
Number of items per task	One
Time allocated for task	Preparation time: 40 seconds Response time: 50 seconds
Instructions for candidates (task rubric)	You are going to give a summary. First read and listen to the task. You can make notes while you listen. Your notes will not be marked. Use the 'Notes' button below to open your notepad. You will then have some time to think about what you want to say. The clock shows how much time you have to give your summary. Start speaking when you hear the tone. [Task details (see 'Question format and presentation' and 'Description' below)]
Question format and presentation	Your tutor has asked you to summarize some research for your tutor group. Listen to two experts talking about [topic]. The two experts make the same two main points. You should: • combine the information from the two experts and • summarize the two main points the experts make. Now listen to the two experts. • Expert 1 • Expert 2 You now have some time to think about what you want to say.
Targeted CEFR level(s)	B2–C1
Targeted cognitive functions	☑ Monitoring
	☑ Articulation
	☑ Buffer
	☑ Formulation
	☑ Planning
	☑ Conceptualization

14.4.2 Speaking Part 3 features of the input

Description	Audio-written situation plus 2 audio-only monologues. The domain will be either academic or professional. Academic monologue text types will be lecture extracts, academic light presentations (such as TEDx-type talks). Professional monologues will be business light presentation, business-related TEDx-type talks. The first monologue establishes the topic. Each monologue contains the same two main points, each with one or more supporting details.
Word count	The total word count for the audio monologues is 220–245 words, a maximum of 125 words for an individual monologue.
Grammatical and/or phonological features	• At C1 level, a range of grammatical and phonological features can be expected. • Main points and supporting details will not be excessively signposted in the monologues (e.g. 'the first main point is…'), although some discourse markers to signal the structure of the talk can occur as these will make the talk sound more natural. • The structure of the monologues will not give away the main points. Instead, main points can come before or after supporting details. • Contractions that occur in everyday life are expected in the monologues, as are features of connected speech e.g. assimilation, elision and linking.
Content (specificity, neutrality, concreteness)	Concrete/specific
Lexical range (refer to Oxford 5000 for CEFR levels of individual words)	Primarily A1–B2 lexis. Questions are written to be accessible to B2 level test takers but allow C1 test takers to demonstrate their proficiency.
Genre/domain (e.g. personal/professional/education/other)	Academic/professional

14.4.3 Speaking Part 3 features of the expected response

Description	Test takers are expected to effectively summarize the main points of the monologues, and part of the task requirement is the ability to distinguish between the main and the supporting points. Test takers are scored on their ability to extract and present the key information in a coherent and intelligible form. The response will be approximately 90–100 words in length. Test takers should aim to use their own words where possible, avoiding using too much language from the task input. Test takers are able to make notes in the on-screen text box.
Presentation (written/recorded/illustrations/options)	Verbal; audio-recorded. After the monologues are heard, test takers are allowed 40 seconds to prepare. A timer is visible counting down from 40 seconds. Recording automatically starts following the preparation. It ceases after 50 seconds once the timer reaches zero seconds. The task is not intended to be a memory exercise. In order to extract the main points from lectures, test takers are able to take notes on screen as there is no textual reference beyond

	the visual prompt. To make notes while listening to the experts, the test taker clicks the Notes button on the ribbon at the bottom of the screen.
Discourse functions	☑ Organizing and sustaining extended discourse (50 seconds)
	☑ Responding and expressing oneself appropriately
	☑ Using appropriate genre conventions
Scoring model	Part 3 is marked against both the standard marking criteria (see Appendix 5) and against a task-specific marking guide. The task-specific marking guide sets out content points which must be included in the response. Part 3 is marked holistically alongside Parts 1 and 2 (as part of Script 1). Test takers receive four scores. Scores for 'Task fulfilment', 'Pronunciation and fluency', 'Grammar' and 'Lexis'. These are added together with scores from Script 2 (Parts 4 and 5). The total score is then translated into a whole CEFR level.

14.4.4 Speaking Part 3 sample task

This task is an extract from an *Oxford Test of English Advanced* practice test. To see how the task displays on screen in the test, see the next section.

Summary

You are going to give a summary. First read and listen to the task. You can make notes while you listen. Your notes will not be marked. Use the 'Notes' button below to open your notepad.

You will then have some time to think about what you want to say. The clock shows how much time you have to give your summary. Start speaking when you hear the tone.

Your tutor has asked you to summarize some research for your tutor group.

Listen to two experts talking about research into breathing. The two experts make **the same two main points.**

You should:
• combine the information from the two experts and
• summarize the **two main points** the experts make.

Now listen to the two experts.

• Expert 1
• Expert 2

You now have some time to think about what you want to say. You can make notes if you wish.

14.4.4.1 Audioscript

Expert 1: Today we are going to talk about some of the claims being made about the benefits of breathing. One of the most common ideas, consistently being put forward, is that learning to breathe better enhances quality of life and reduces stress levels. Focusing on breathing and training ourselves to breathe in certain ways has a respectable history in traditions such as yoga and meditation, and significant evidence exists to link good breathing with maintaining good mental and physical health. Another, more far-reaching claim is that certain breathing techniques can actually cure some illnesses, though more solid evidence from high-quality scientific research is needed to back these up.

Expert 2: The benefits of breathing have been discussed for centuries. A growing number of studies have confirmed that breathing techniques can help to manage issues such as stress and support general good health. By following certain practices, such as breath focus and equal breathing, not only can we relieve stress and feel calmer but also reduce our chances of becoming ill. In more recent years, suggestions have been made that breathing techniques can be used to successfully treat and heal specific medical conditions. Whilst these findings are promising, no firm conclusions can be drawn without further research.

14.4.4.2 Task-specific marking guide

Main point	E.g. mention or paraphrase of one of the following
Main point 1: There is evidence that breathing well can help people maintain good health	Evidence shows breathing better improves quality of life/reduces stress/supports general good health
Main point 2: More research is needed into claims that breathing well can cure physical illness that people already have	There are claims that breathing techniques can cure/treat/heal some medical conditions/illnesses
	No firm conclusions/more research needed to back these up

14.4.5 Speaking Part 3 screen display

14.5 Speaking Part 4 (Debate) design and sample tasks

Module	Speaking
Part	4
Task label	Debate
CEFR	B2–C1

14.5.1 Speaking Part 4 task features

Skill focus	- Arguing a case on a complex issue - Structuring an argument - Formulating points precisely - Employing emphasis effectively
Description	Putting a case for or against a proposition in a debate. The aim of the debate is to elicit the spoken language features required to give an extended turn.
Number of items per task	One
Time allocated for task	Preparation time: 45 seconds Response time: 2 minutes
Instructions for candidates (task rubric)	Your tutor has asked you to take part in a class debate. You are going to put a case for or against the following statement: ['insert statement.'] Prepare your case for the debate. You should: • use two OR three of the ideas below to argue your case • provide support for the ideas you choose • give a conclusion. You now have some time to prepare. You can make notes if you wish.
Question format and presentation	Test takers are presented with an audio-written standard rubric. The task and the debate statement are audio-written, supported by the (screen-only) visual diagram, plus the time remaining for the statement shown in an onscreen clock. The visual repeats the debate statement, which is in the centre. Around the debate statement are five prompts, written in the form of noun phrases.
Targeted CEFR level(s)	B2–C1
Targeted cognitive functions	☑ Monitoring
	☑ Articulation
	☑ Buffer
	☑ Formulation
	☑ Planning
	☑ Conceptualization

14.5.2 Speaking Part 4 features of the input

Description	In Debate tasks, the debate statement requires the test taker to take a position on whether a proposition is right. Prompts should be as neutral and objective as possible so that they can be used whether the test taker adopts a position for or against the debate statement. The Debate tasks are written to be accessible to target test takers. They are constructed to elicit a range of spoken language features and provide scope for test takers to expand on each of the ideas in the prompts, drawing on their personal experience and general knowledge.
Word count	Standard rubric (see above) Debate statement: 15 words maximum Visual: Debate statement: 15 words maximum

	Individual prompts: 5 words maximum individually (however, maximum 20 words for all prompts)
Grammatical and/or phonological features	The Debate tasks should be accessible to the target test takers. The tasks should be constructed to elicit a range of spoken language features and provide scope for test takers to expand on each of the ideas in the prompts, drawing on their personal experience and general knowledge. The Debate is written with Part 5 (Follow-up questions) in mind – the topic in the Debate lends itself to complex evaluative questions in Part 5.
Content (specificity, neutrality, concreteness)	Concrete/specific
Lexical range (refer to Oxford 5000 for CEFR levels of individual words)	Primarily A1–B2 lexis. The language of the rubric should be accessible to test takers at B2 level but to allow C1 test takers to demonstrate their proficiency.
Genre/domain (e.g. personal/professional/education/other)	Professional/academic

14.5.3 Speaking Part 4 features of the expected response

Description	Test takers are expected take a position and to address two or three of the prompts, sustaining their argument by addressing each prompt at some length, providing salient examples and employing appropriate rhetorical devices to persuade the listener. The turn is expected to be well organised, with clear signalling of the overall structure of the turn, and suitable signposting as they move from one prompt to the next, rounding off with a suitable ending.
Presentation (written/recorded/illustrations/options)	Verbal; audio-recorded. The test taker should aim to give a response of approximately 250–300 words. After the task rubric is heard, the timer is visible, counting down from 45 seconds. Recording automatically commences once the preparation time ends. Recording automatically ceases once timer reaches zero seconds. To make notes, the test taker clicks the Notes button on the ribbon at the bottom of the screen.
Discourse functions	☑ Organizing and sustaining extended discourse
	☑ Developing and sustaining an argument
	☑ Giving opinions on complex issues
	☑ Highlighting significant points appropriately
Scoring model	Part 4 is marked holistically alongside Part 5 (as part of Script 2). Test takers receive four scores. Marks are awarded for 'Task fulfilment', 'Pronunciation and fluency', 'Grammar' and 'Lexis'. These are added together with scores from Script 2 (Parts 4 and 5). The total marks are then translated into a standardized score and CEFR level.
Supplementary information/criteria	The prompts are written to be in some way contentious.

14.5.4 Speaking Part 4 sample task

This task is an extract from an *Oxford Test of English Advanced* practice test. To see how the task displays on screen in the test, see the next section.

Debate

You are going to take part in a debate. First read and listen to the task, then decide what you want to say. The clock shows how much time you have to speak. Start speaking when you hear the tone.

Your tutor has asked you to take part in a class debate. You are going to put a case for or against the following statement:

'Cheap, mass-produced clothes should be banned.'

Prepare your case for the debate. You should:

- use **two OR three** of the ideas below to argue your case
- provide support for the ideas you choose
- give a conclusion.

You now have some time to prepare. You can make notes if you wish.

	impact on the environment	
effect on shops		quality of clothing
low prices	Cheap, mass-produced clothes should be banned.	employment issues
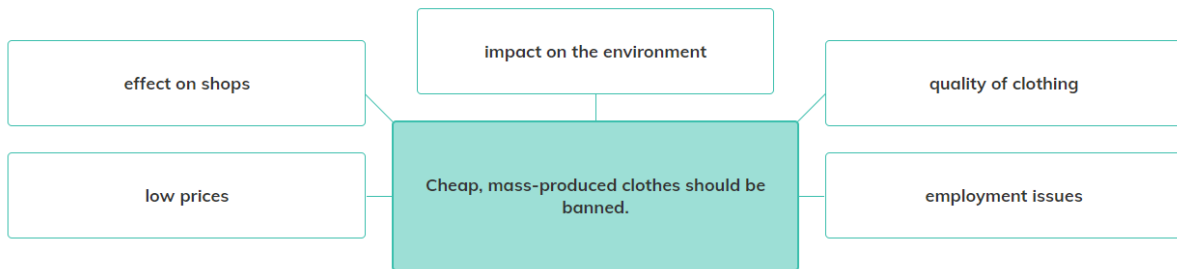
14.5.5 Speaking Part 4 screen display

14.6 Speaking Part 5 (Follow-up questions) design and sample tasks

Module	Speaking
Part	5
Task label	Follow-up questions
CEFR	B2–C1

14.6.1 Speaking Part 5 task features

Skill focus	- Responding to questions - Responding and expressing opinions appropriately - Reinforcing key messages
Description	Four audio-only questions, thematically related to the Debate (Part 4). The aim of Part 5 is to elicit the more complex spoken language features required to develop discourse on general issues (rather than the more personal, everyday topics targeted in Part 1 – Interview).
Number of items per task	Four thematically related questions.
Time allocated for task	Response time: 40 seconds per question (2 minutes 40 seconds total).
Instructions for candidates (task rubric)	You are going to answer four questions on the topic of your debate. The clock shows how much time you have to speak. Start speaking when you hear the tone.
Question format and presentation	Test takers are presented with an audio-written standard rubric. The task questions are audio-only, supported by the (written-only) number of the question appearing on the screen (e.g. 'Question 1') for the duration of the time allotted for the question being asked, plus the time remaining for the question shown in an onscreen clock. Question numbers do not form part of the audio.
Targeted CEFR level(s)	B2–C1
Targeted cognitive functions	☑ Monitoring
	☑ Articulation
	☑ Buffer
	☑ Formulation
	☑ Planning
	☑ Conceptualization

14.6.2 Speaking Part 5 features of input

Description	- Question 1 in Part 5 directly relates to the theme in the debate, while questions 2–4 move progressively away from the debate theme in order to elicit a wide range of spoken language features, and to avoid topic overlap amongst the questions. - Test takers are presented with an **audio-written** standard rubric. The task questions are audio-only, supported by the (written-only) number of the question appearing on the screen (e.g. 'Question 1') for the duration of the time allotted for the question to be asked, plus the time remaining for the question shown in an onscreen clock. Note that the question numbers do not form part of the audio.

Word count	**25 words** per question (maximum).
Grammatical and/or phonological features	- Questions are direct (single or multi-clause) but may have more than one sentence. The first sentence can establish a proposition, followed by a question asking the test taker's opinion. - Questions avoid all but the most basic contractions. Questions may be fronted with buffers (e.g. 'Can you tell me what you have done with your friends recently.', **NOT** 'What have you done with your friends recently?'). - Questions may ask for a direct opinion or be scenario-style questions. These establish hypothetical situations and ask test takers to respond.
Content (specificity, neutrality, concreteness)	Concrete/specific/abstract
Lexical range (refer to Oxford 5000 for CEFR levels of individual words)	Primarily A1–B2 lexis. Questions are written to be accessible to B2 level test takers but allow C1 test takers to demonstrate their proficiency.
Genre/domain (e.g. personal/professional/education/other)	Personal/professional/academic

14.6.3 Speaking Part 5 features of expected response

Description	The task is designed to elicit responses around the test taker's opinion, and for them to draw upon their world knowledge.
Presentation (written/recorded/illustrations/options)	Verbal; audio-recorded. After question is heard, audio recording commences immediately. A timer is visible, counting down from 40 seconds. Recording automatically ceases once timer reaches zero seconds.
Discourse functions	Giving descriptions
	Giving information
	Comparing and contrasting
	Agreeing and disagreeing
	Expressing, justifying and responding to opinions
	Expressing feelings
	Speculating
	Making critical remarks
	Expressing disagreement diplomatically
	Highlighting and categorizing main points
	Presenting ideas in a logically connected way.
Scoring model	Part 5 is marked holistically alongside Part 4 (as part of Script 2). Test takers receive four scores. Scores for 'Task fulfilment', 'Pronunciation and fluency', 'Grammar' and 'Lexis'. These are added together with scores from Script 2 (Parts 4 and 5). The total score is then translated into a whole CEFR level.
Supplementary information/criteria	Questions should be straightforward but elicit a range of possible language structures.
Example question frames (this is not exhaustive) - Some people say ……………. What's your view/opinion? - How important is it to …………………………? - How far do you agree that ………………………? - Do you agree that ……………? (Why / Why not?)	

- One day ……………… How likely do you think this is?
- What are the advantages/disadvantages of …………..?
- What are the most significant aspects of………………………?
- What are the issues surrounding………………………?

Example of scenario-style question frames (this list is not exhaustive)
- If you were ………………………, what would you do about………………………?
- Imagine you were ………………………. What actions would you take in this scenario?

14.6.4 Speaking Part 5 sample tasks

This task is an extract from an *Oxford Test of English Advanced* practice test. To see how the task displays on screen in the test, see the next section.

Follow-up questions
You are going to answer four questions on the topic of your debate.
The clock shows how much time you have to speak.
Start speaking when you hear the tone.

1. Your talk was about fashion. Do you think it's possible to tell what someone is like as a person from the clothes they wear?
2. Many young people say there's too much pressure to be fashionable. What do you think?
3. Do you agree that clothes are important in giving people a sense of belonging to a group?
4. Some people say that it doesn't matter what you wear, it's more important how you behave. What do you think?

14.6.5 Speaking Part 5 screen display

Individual questions are heard but not displayed on the screen.
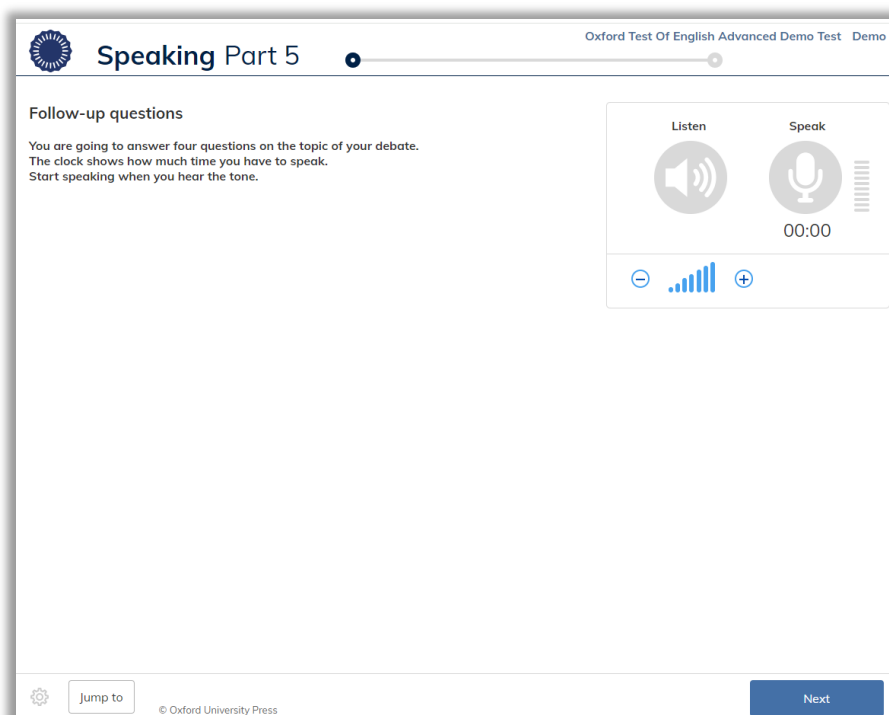
15 Appendix 2: Listening module details

15.1 Listening module construct and domain definitions

Test purpose	The *Oxford Test of English Advanced* Listening module assesses test takers' aural English proficiency for three sub-domains: academic, professional, and social purposes. Within these sub-domains, the test assesses the ability to: • identify clearly stated information • identify implicit meaning, opinion, attitude and speaker's purpose in a monologue or dialogue • identify structural and organizational features • listen for gist • identify the development of ideas.			
Examples of domain activities	Following a lecture (live; recorded), following a seminar discussion or a debate; content will be academic in nature (i.e. topics of general academic interest but not dependent on specific academic vocabulary. Accents; English as a lingua franca (ELF). Speakers using general English with standard grammar/spoken grammar, but with a variety of regional accents.			
Communication goals by sub-domain		Academic	Professional	Social
	To identify information, ideas and opinions, agreement and disagreement	☑	☑	☑
	To identify implicit meaning in a range of audioscripts	☑	☑	☑
	To identify detail and implied meaning	☑	☑	☑
	To identify the speaker's purpose	☑	☑	☑
	To identify shifts in emphasis/focus	☑	☑	☐
	To identify register shifts	☑	☑	☐
	To identify discourse organization	☑	☑	☐
Targeted cognitive functions	**Decoding:** the test taker analyzes aural information to identify individual phonetic segments and syllables.			
	Lexical search: once specific word or phrasal forms have been identified, the test taker matches the form to known words using their linguistic resources such as vocabulary (lexicalization).			
	Parsing: the test taker imposes a syntactic pattern on a group of words to form propositions.			
	Meaning construction: the test taker links propositions together in their mind so that they build a mental model of what is heard.			
	Discourse construction: as the test taker progresses, they integrate new information into what has been heard so far, relating the content and meaning of the audio to their own knowledge to provide context.			

15.2 Listening Parts 1 and 4 (Short dialogues and monologues) design and sample tasks

Module	Listening
Part	1 and 4
Task label	Short dialogues and monologues
CEFR	B2 or C1

15.2.1 Listening Parts 1 and 4 task features

Skill focus	- Identifying specific information - Understanding abstract ideas - Understanding implied meaning
Description	Short monologues/dialogues Each monologue/dialogue has one 3-option multiple-choice question (text options, or picture options for B2-level tasks)
Number of items per task	Five discrete 3-option MCQs
Audioscript	For each item: Monologue or Dialogue between male and female speakers.
Instructions for candidates (task rubric)	Listen and choose the correct option.
Question format and presentation	Each item consists of an instruction, a stem which can be either a question (e.g. 'What does the professor claim about X?') or an unfinished sentence (e.g. 'The professor argues that…') followed by three options.
Description of the item type (number of words/lexical range – refer to Oxford 5000/grammar/illustrations)	Situation (optional): 10 words maximum Stem: 10 words maximum Options: 10 words each
Targeted CEFR level(s)	B2 or C1
Targeted cognitive functions	☑ Discourse construction ☑ Meaning construction ☑ Parsing ☑ Lexical search ☑ Input decoding

15.2.2 Listening Parts 1 and 4 features of the audioscript

	B2	C1
Word count	70–96 words per audioscript. Word counts for the audioscript do not include the identifiers for the speakers i.e. M/F.	120–150 words per audioscript. Word counts for the audioscript do not include the identifiers for the speakers i.e. M/F.
Discourse features	- turn-based interactions - minimal idiomatic expressions and colloquialisms - explicit register shifts - implication - speakers reveal attitude, mood and intentions - minimal phonetic reduction (segments, syllables, or words are uttered with short duration and/or less articulation, or omitted entirely, to mimic natural speech) - no overlapping turns	- complex interactions - idiomatic expressions and colloquialisms - register shifts - implication - speakers reveal attitude, mood and intentions - phonetic reduction (segments, syllables, or words are uttered with short duration and/or less articulation, or omitted entirely, to mimic natural speech) - overlapping turns

Lexical range (refer to Oxford 5000)	Primarily A1–B2. C1 lexis permitted, although is not tested, as defined in the Oxford 5000.	Primarily A1–C1. C2 lexis permitted, although is not tested, as defined in the Oxford 5000.
Range of grammatical features present/allowed	Grammatical features appropriate for the level.	No restrictions on range of grammatical features permitted.
Content (specificity, neutrality, concreteness)	Concrete and abstract, including implied meaning.	Concrete and abstract, but with a strong emphasis on implied meaning.
Genre/domain (e.g. personal/professional/education/other)	Social/academic/professional	Academic/professional. At C1 level, topics will be more academic or vocational in nature and more abstract in comparison to lower levels. Topics include general academic interest in science, humanities, business and sports accessible to non-specialist audiences. Conversations might be between a professor/lecturer/tutor and a student about an upcoming assignment.

15.2.3 Listening Parts 1 and 4 features of the expected response

Description	Selection of a single option from three available.
Presentation (written/recorded/illustrations/options)	On-screen, with text options vertically arranged, or picture options horizontally arranged.
Scoring model	Independent items. Computer-adaptive scoring model with algorithm which selects the subsequent item depending on the response to the preceding item. All items contribute to each new estimate of test takers' ability. The updated ability after the 5th task is used to select the Part 2 task.
Supplementary information/criteria	At C1, audioscripts will contain some of the following: • overlapping turns • longer individual turns in a dialogue in comparison to lower CEFR levels • digressions • false starts
Example question types	**C1 question types** **Task type 1: Gist** These tasks target test takers' understanding of the gist of the conversation/presentation. • e.g. What is the talk mainly about? • e.g. What is the presenter going to explain? **Task type 2: Function / Reason / Purpose**

These tasks target test takers' understanding of the function of / reason for / purpose of the conversation. The key contains the information that is implicitly stated across the audioscript.

- e.g. Why does the student go to see the professor?
- e.g. Why does the professor give the example of X?

Task type 3: Feelings / Opinions / Attitude

These tasks target test takers understanding of the feelings / opinions / attitude of one of the speakers in the conversation. The key contains the information that is implicitly stated across the audioscript.

- e.g. What is the person's attitude/opinion of … ?
- e.g. What can be inferred about X's attitude towards X?

Task type 4: Type

These tasks require the test taker to identify the genre/type of the talk.

- e.g. What kind of talk was given?
- e.g. Where would this talk most likely take place?

Task type 5: Topic

These tasks are designed to target the topic focus of the speakers.

- e.g. What topic are the student and the professor talking about?
- e.g. What is the main topic of the lecture?

15.2.4 Listening Parts 1 and 4 sample task – C1 level

This task is an extract from an *Oxford Test of English Advanced* practice test. To see how the task displays on screen in the test, see the next section.

Listen and choose the correct answer.

Two journalists are discussing gender stereotypes in the media.
In which area does the man feel there has been progress?

A the way men and women are portrayed in advertising
B the roles played by girls and boys in children's TV
C the proportion of male and female participants in debates

Audio script:

---

4 Two journalists are discussing gender stereotypes in the media. In which area does the man feel there has been progress?

A the way men and women are portrayed in advertising

B the roles played by girls and boys in children's TV

C the proportion of male and female participants in debates

F I listened to a radio discussion today on the environment, and do you know what… all the experts invited to contribute were men.

M I heard it, too. And all it would've taken is for the programme editor to pick up the phone to a few universities, to put a more balanced panel together. There's no excuse for it. Though it's not all doom and gloom … Take commercials – remember when washing up was only done by women and cars were only driven by men?

F And kids' TV has come a long way too, wouldn't you say?

M I'd like to think so – the days when boys were routinely cast as leaders and girls as helpers should be long gone – though it seems we're still exposing kids to this type of thing. The producers need to be called out!

F The more pressure the better, I'd say!

Now listen again.

Answer key: A

15.2.5 Listening Part 1 screen display

15.3 Listening Part 2 (Note completion) design and sample tasks

Module	Listening
Part	2
Task label	Note completion
CEFR	B2 or C1

15.3.1 Listening Part 2 task features

Skill focus	Listening to identify specific information	
	B2	**C1**
Description	A longer informational/descriptive monologue with note-completion task (3-option MCQ). Test takers complete notes summarizing the key information.	A longer informational/descriptive monologue with note-completion task (open cloze). Test takers complete notes summarizing the key information.
Number of items per task	Five 3-option MCQs plus one example item.	Open cloze with six items (maximum of two words per gap).
Audioscript	One monologue, in which the speaker has useful information to impart to the listener.	
Instructions for candidates (task rubric)	Complete the notes with a word or two-word phrase from the audio. Remember to check your spelling. The clock shows how much time you have to look at the task.	
Question format and presentation	Input is presented to test takers in the form of incomplete Notes. Notes are separated beneath subheadings. Notes are complete sentences with words removed. Subheadings are not gapped. Gaps can come at the start or end of a sentence but will normally come in the middle of a sentence. Test takers are able to type in their responses to see the response in context.	
Description of the item type (number of words/lexical range – refer to Oxford 5000/grammar/illustrations)	Instructions for candidates: 12 words maximum, including 'listen to'. Notes: 50–100 words (including gaps). The word count will be evenly spread across the items.	
Targeted CEFR level(s)	B2 or C1	
Targeted cognitive functions	☑ Discourse construction	
	☑ Meaning construction	
	☑ Parsing	
	☑ Lexical search	
	☑ Input decoding	

15.3.2 Listening Part 2 features of the audioscript

	B2	**C1**
Word count	350–450 words per audioscript.	450–550 words per audioscript.
Discourse features	- organizational features of a talk (introduction, body, conclusion) - speaker talking directly to audience - no specialist vocabulary	- lecture conventions (introduction, body, conclusion) - speaker talking directly to audience

	• more explicit verbal cues for shift in emphasis/focus • more explicit signposting	• little to no specialist vocabulary • verbal cues for shift in emphasis/focus • relevant signposting
Lexical range (refer to Oxford 5000)	Primarily A1–B2. C1 lexis permitted, although is not tested, as defined in the Oxford 5000.	Primarily A1–C1. C2 lexis permitted, although is not tested. Key words and distractors will be at CEFR level B2 as defined in the Oxford 5000.
Range of grammatical features present/allowed	Grammatical features appropriate for the level.	No restrictions on range of grammatical features permitted.
Content (specificity, neutrality, concreteness)	Concrete and abstract, including implied meaning.	Concrete and abstract, but with a strong emphasis on implied meaning.
Genre/domain (e.g. personal/professional/education/other)	Social/professional/academic	Professional/academic. At C1 level, topics will be more academic or vocational in nature and more abstract in comparison to lower levels. Topics include general academic interest in science, humanities, business and sports accessible to non-specialist audiences. The talk should be a believable example of one which lends itself to a note-taking activity in which the listener needs to note down the most important details from a talk.

15.3.3 Listening Part 2 features of the expected response

	B2	C1
Description	Selecting options from 3-option MCQs.	Typing a response (1–2 words) into text boxes for each item.
Presentation (written/recorded/illustrations/options)	On-screen, notes arranged under relevant subheadings. The text around the gap paraphrases audioscript content. Each item is contained within a single sentence/paragraph. Items (in the Notes) come in the same order as information in the audioscript.	
Scoring model	The items are grouped. Each item has a difficulty value, and the task has an overall average difficulty value. The selection algorithm selects tasks based on their average difficulty. The majority of items are at the same CEFR level as the overall average difficulty. For example, for a C1 task, a minimum of four out of six items will be C1; for a B2 task, a minimum of three out of five items will be B2. Items	

	falling outside the overall difficulty average must be within .3 of a level of the overall difficulty average. All items contribute to new estimate of test takers' ability. The updated ability is used to select Part 3 task. The algorithm will select a task with a different topic to that presented in Part 2.
Supplementary information/criteria	For C1 tasks: - The correct response for each item will be verbatim from the audioscript. - Correct responses will be a maximum of two words. - Correct responses will not be jargon or technical words. - American or British spellings of correct responses are acceptable. - Gaps target noun phrases in the audioscript, not verbs or adjectives. - Correct/incorrect responses will not be dependent on plural/singular forms of noun phrases. - Correct responses will not appear anywhere else in the Notes.

15.3.4 Listening Part 2 sample task – C1 level

This task is an extract from an *Oxford Test of English Advanced* practice test. To see how the task displays on screen in the test, see the next section.

> **Listen to a lecture about a seed bank called the Svalbard Global Seed Vault. Complete the gaps in the notes with a word or two-word phrase from the audio. Remember to check your spelling.**
>
> **The clock shows how much time you have to look at the task.**

The Svalbard Global Seed Vault

History

The new seed bank replaced one that was housed in an old [1] .... on Svalbard.

The facility

The current facility was built in rock to avoid the [2] .... sometimes found in soil.

There is an impressive [3] .... outside the seed bank.

Between 2016 and 2019, the building's [4] .... had to be replaced.

The seeds

Seeds can only be removed from the bank by the original owner.

Seeds are stored at -18°C in specially designed [5] .... made from metal.

Conclusion

Seeds that can withstand severe conditions may help with the consequences of [6] ....

Audioscript:

Welcome to my lecture on the Svalbard Global Seed Vault. As its name suggests, this is a bank of currently well over one million seeds, located on the Norwegian island of Svalbard. It was established in 2008, although a smaller, national seed bank had existed there since the 1980s, a stone's throw from the island's airport. Its coal mine location was increasingly being seen as unsuitable, so discussions began about not only rehousing the seed bank, but creating something altogether more ambitious.

What transpired was the foundation of the Svalbard Global Seed Vault, whose remit was to accept duplicates of seeds from national and regional seed banks around the world, thereby safeguarding the future of countless plant species. The earth in Svalbard is permanently frozen, so there's little available moisture which could damage seeds. However, dangerous gases occasionally build up in the cavities, so it was decided to create the bank within the rock of a nearby hillside.

From the outside, there are few clues as to its function apart from a modest vault door in a tall concrete shaft, which slopes gradually into the ground. For something as critical as this facility, one might envisage there being a range of security features surrounding it. The reality is somewhat different – an art installation at the top, made of steel, mirrors and prisms reflecting the Arctic lights, which hints at its being something of true significance.

The interior of the facility boasts three huge halls, hidden a hundred metres inside the hill. Each is capable of holding one and a half million seeds. Extensive renovations were carried out between 2016 and 2019, during which the main entrance to the facility was made substantially more waterproof. Also, an environmentally friendly cooling system was installed – the original one being unable to adequately regulate the humidity.

There are three scheduled deposits of seeds per year at the bank. Ownership of the seeds remains with whichever individual or national institution placed them in the bank. To prevent the precious seeds falling into the wrong hands, the security team ensure that black box procedures are followed, meaning that no-one other than the original owner is allowed to withdraw seeds from the bank.

The seeds are kept at a decidedly chilly minus eighteen degrees Celsius. In order to prevent any spoilage, it was crucial that anything which came into direct contact with the seeds should be unreactive, so the custom-made envelopes they're kept in are fabricated from aluminium. These are then placed in plastic crates, which are stacked on the shelves in the facility's halls.

So in summary, the Svalbard Global Seed Vault acts as invaluable insurance against any of the world's many plant and crop species going extinct, and for maintaining genetic diversity. It's there to cover relatively low-impact eventualities, such as local mismanagement of plant resources. But its role in potentially combatting the effects of climate change is becoming increasingly apparent. After all, there are seeds within it from species that are capable of coping with the most extreme environments on earth, from burning deserts to frozen tundra.

Now listen again.

Answer key:

1	coal mine
2	dangerous gases
3	art installation
4	cooling system
5	envelopes
6	climate change

15.3.5 Listening Part 2 screen display

15.4 Listening Part 3 (Matching) design and sample tasks

Module	Listening
Part	3
Task label	Matching
CEFR	B2 or C1

15.4.1 Listening Part 3 task features

Skill focus	• Identifying stated opinions • Identifying implied meaning
Description	A dialogue between a man and a woman. Items focus on opinions that feature in the text. Test takers have to identify who expressed the opinion: the woman, the man or both.
Number of items per task	B2: Five 3-option multiple-choice questions matching the opinion to the speaker – woman, man, or both. C1: Six 3-option multiple-choice questions matching the opinion to the speaker – woman, man, or both.

Audioscript	One dialogue between a male and a female
Instructions for candidates (task rubric)	Listen to [description of speakers > situation > topic]. Match the people (the woman, the man, or both) to the opinions below. The clock shows how much time you have to look at the task.
Question format and presentation	Questions are presented on-screen, vertically. Each item is an affirmative statement, representing an opinion stated in the audioscript. Three options (woman, man, both) are presented for each item, next to the item stem.
Description of the item type (number of words/lexical range – refer to Oxford 5000/grammar/illustrations)	Statements: 16 words maximum.
Targeted CEFR level(s)	B2 or C1
Targeted cognitive functions	☑ Discourse construction
	☑ Meaning construction
	☑ Parsing
	☑ Lexical search
	☑ Input decoding

15.4.2 Listening Part 3 features of the audioscript

	B2	C1
Word count	400–525 words	600–650 words
Discourse features	- no specialist vocabulary - explicit verbal cues for shift in emphasis/focus - more explicit signposting for agreement/disagreement - grammar and syntax level appropriate - minimal idiomatic expressions and colloquialisms - explicit register shifts - implication - opinion and attitude - reduced phonetic reduction (segments, syllables, or words are uttered with short duration and/or less articulation, or omitted entirely, to mimic natural speech)	- little to no specialist vocabulary - verbal cues for shift in emphasis/focus - signposting for agreement/disagreement - complex sentences with high-level syntax - idiomatic expressions and colloquialisms - register shifts - implication - opinion and attitude - phonetic reduction (segments, syllables, or words are uttered with short duration and/or less articulation, or omitted entirely, to mimic natural speech) - redundancy - overlapping turns
Lexical range (refer to Oxford 5000)	Primarily A1–B2. C1 lexis permitted, although is not tested, as defined in the Oxford 5000.	Primarily A1–C1. C2 lexis permitted, although is not tested, as defined in the Oxford 5000.

Range of grammatical features present/allowed	Grammatical features appropriate for the level.	No restrictions on range of grammatical features permitted.
Content (specificity, neutrality, concreteness)	Concrete and abstract	Concrete and abstract
Genre/domain (e.g. personal/professional/education/other)	Social/general interest	Academic/professional. At C1 level, topics will be more academic or vocational in nature and more abstract in comparison to lower levels. Topics include general academic interest in a variety of topics, accessible to non-specialist audiences. In order to be able to speak about the abstract topics with the appropriate high-level lexis, the speakers will normally be from that field. The expert commentators will still focus on opinions and attitudes in relation to the topic.

15.4.3 Listening Part 3 features of the expected response

Description	Selection of a single option from three available.
Presentation (written/recorded/illustrations/options)	On-screen, vertically arranged.
Scoring model	The items are grouped. Each item has a difficulty value, and the task has an overall average difficulty value. The selection algorithm selects tasks based on their average difficulty. All items contribute to new estimate of test takers' ability. The updated ability is used to select the first Part 4 task. The algorithm will select a task with a different topic to that presented in Part 3.
Supplementary information/criteria	Items will be independent, with no overlap in textual references between items.

15.4.4 Listening Part 3 sample task – C1 level

This task is an extract from an *Oxford Test of English Advanced* practice test. To see how the task displays on screen in the test, see the next section.

> **Listen to two fashion journalists discussing clothing made from recycled plastic.**
> **Match the people (the woman, the man, or both) to the opinions below.**

	Woman	Man	Both
1 Some brands of clothing made from recycled plastic have become extremely popular.	◯	◯	◯
2 The range of sources of plastic for recycling into clothing is very wide.	◯	◯	◯
3 There are some environmental drawbacks to recycling plastic into clothing.	◯	◯	◯
4 Recycled-plastic clothing should come with information on how it was made.	◯	◯	◯
5 Clothing made from recycled plastic is as comfortable as other synthetic clothing.	◯	◯	◯
6 There should be no tax on clothing made from recycled plastic.	◯	◯	◯

Audio script:

Listen to two fashion journalists discussing clothing made from recycled plastic. Match the people (the woman, the man, or both) to the opinions below. You now have 30 seconds to look at the task.

M Hi, I'm fashion journalist Ronnie Callow, and I'm delighted to welcome fellow writer Elena Bartlett to this week's podcast, in which we'll be discussing clothing made from recycled plastic. Elena, a few of the groundbreaking companies who have built their identities around this kind of clothing generate so much excitement on social media …

F I'd say there was certainly an astonishing amount of initial interest when they first appeared on the fashion scene, and all sorts of predictions that they'd take the world by storm. These haven't really materialized, though, for one reason or another, and it's something of a rarity to see celebrities or even people in the street wearing them.

M Well, whatever the fact of the matter is, clothing's certainly being made from some unusual things these days – old plastic fishing nets, for example.

F For sure, and plastic bottles are commonly recycled and repurposed into synthetic fabrics. I guess the sky's the limit in terms of what you can transform into clothing …

M Actually, my understanding is that the type and grade of plastic required is harder to come by than people perhaps imagine, so diverting the entire contents of our recycling bins to create materials suitable for clothing is unlikely to be happening anytime soon.

F We'll see, but what's for sure is that the processes involved in transforming plastic into fabric aren't straightforward, so …

M This is what troubles me somewhat about recycling plastic into clothing. I wonder whether it gives the rather misleading impression that manufacturers and retailers are creating garments in a sustainable way …

F Whereas in fact they're just cutting down slightly on the resources required, but certainly not the energy used, particularly when compared to producing items of clothing from brand new synthetic fabrics.

M So maybe it's a step in the right direction, maybe not.

F I've heard it suggested that clothing manufacturers who used recycled plastic in their products should be required to state details about exactly how garments were made on the product's label.

M It'd mean that companies trying to do their bit environmentally would be faced with complying with extra legislation, while the ones who couldn't care less about the environment could get away with doing next to nothing.

F It'd encourage being transparent about the impact manufacturing has environmentally, though, and other companies might well follow. Anyway, do you own any clothes made from recycled plastic?

M I do – some swimming trunks and a jacket – and I've tried on plenty of others. In terms of how they feel to wear, it's actually tremendously hard, if not impossible, to tell them apart from garments created from artificial fibres.

F I've run testing sessions recently, in which participants have been given a coat, for example, made from recycled plastic, and a similar one that hadn't been recycled in any way. As far as cosiness went, there was next to no difference between the scores they gave for each – apart from anything made from natural fibres, which did tend to outperform them.

M Fascinating! Of course one way to encourage more clothing companies to use recycled fabrics is to remove any government tariffs or other charges from products made from them.

F I'd much rather that approach was applied to products made from natural fibres which have had minimal processing during their manufacture instead. Otherwise you might end up sending out the wrong message.

M Every little helps, though, surely, so I can see no reason for going down that road for both of these kinds of garments. Anyway, it's been really great talking to you, Elena – thanks so much for joining me.

F You're welcome!

Now listen again.

Answer key:

1	Man
2	Woman
3	Both
4	Woman
5	Both
6	Man

15.4.5 Listening Part 3 screen display

16 Appendix 3: Reading module details

16.1 Reading module construct and domain definitions

Test purpose	The *Oxford Test of English Advanced* Reading test assesses test takers' reading English proficiency for three sub-domains: social, academic and professional purposes. Within these sub-domains, the test assesses: • **careful local reading** to identify clearly stated information and ideas, implicit meaning, opinion, attitude and writer's purpose • **expeditious search reading** to identify specific information • **expeditious global reading** for gist and to identify structural and organizational features of a text • **careful global reading** to understand the overall message of a text.			
Examples of domain activities	Reading books for leisure; papers; monographs; research reports; lecture notes; summaries of longer texts; abstracts; business reports. Readers will encounter multiple genres of texts on a single topic.			
Communication goals by sub-domain		Academic	Professional	Social
	To identify information, ideas and opinions	☑	☑	☑
	To identify implicit meaning in a range of texts	☑	☑	☑

	To identify detail and implied meaning	☑	☑	☑
	To identify the writer's purpose	☑	☑	☑
	To identify text structure	☑	☑	☐
	To identify organizational features of a text	☑	☑	☐
	To understand the overall message of a text	☑	☑	☐
Targeted cognitive functions	**Decoding:** the test taker analyzes visual information to identify individual letters and morphological variants.			
	Lexical search: once specific word forms have been identified, the test taker matches the forms to known words using their linguistic resources such as vocabulary (lexicalization).			
	Parsing: the test taker imposes a syntactic pattern on a group of words to form propositions.			
	Meaning construction: the test taker links propositions together in their mind so that they build a mental model of the text.			
	Discourse construction: as the test taker progresses, they integrate new information into what has been read so far, relating the content and meaning of the text to their own knowledge to provide context.			

16.2 Reading Part 1 (Short texts) design and sample tasks

Module	Reading
Part	1
Task label	Short texts
CEFR	B2 or C1

16.2.1 Reading Part 1 task features

Skill focus	Careful reading at local and global levels: • identifying main message • identifying writer's purpose • identifying detail and implied meaning
Description	Short texts each with one three-option multiple-choice question (text options); reading to identify information, ideas and opinions, reading to identify implicit meaning in a range of fiction, non-fiction, general and more specialized academic-style texts.
Number of items per task	Six discrete 3-option multiple-choice questions.
Time allocated for task	Test taker has 1 minute 20 seconds to answer each question.
Instructions for candidates (task rubric)	Read the [insert text type] and choose the correct answer.
Question format and presentation	Each item consists of an instruction (see above); a text (e.g. a professional journal extract); a stem, which can be either a question (e.g. 'What does the author claim about X?'), or an unfinished sentence (e.g. 'The author argues that…'); and three options.
Description of the item type (number of words/lexical range – refer to Oxford 3-5000/grammar/illustrations)	Item (stem and three options) = maximum 55 words. Lexis in stem and options should avoid C2 lexis as defined within the Oxford 5000.
Targeted CEFR level(s)	B2 or C1
Targeted cognitive functions	☐ Creating intertextual representation
	☑ Creating a text-level representation
	☑ Building a mental model
	☑ Inferencing
	☑ Establishing propositional meaning
	☑ Syntactic parsing

	☑ Lexical access	
	☐ Word recognition	

16.2.2 Reading Part 1 features of the input

	B2	C1
Word count	60–70 words	100–120 words
Discourse/text types		
Academic journal extract	☐	☑
Advert	☑	☐
Blog entry	☑	☑
Email	☑	☐
Letter	☑	☑
Newspaper editorial	☐	☑
Note	☑	☐
Notice	☑	☐
Professional journal extract	☐	☑
Report extract	☑	☑
Review	☑	☑
Text message	☑	☐
Range of grammatical features present/allowed	Grammatical features appropriate for the level.	No restrictions on range of grammatical features permitted.
Content (specificity, neutrality, concreteness)	Concrete and abstract, including implied meaning.	Concrete and abstract, but with a strong emphasis on implied meaning.
Lexical range (refer to Oxford 5000)	Primarily A1–B2. C1 lexis permitted, although is not tested, as defined in the Oxford 5000.	Primarily A1–C1. C2 lexis permitted, although is not tested, as defined in the Oxford 5000.
Genre/domain (e.g. personal/professional/education/other)	Personal/professional	Personal/literary/professional/ academic

16.2.3 Reading Part 1 features of the expected response

Description	Selection of a single option from three available. Each option is presented in a lozenge. Test takers can click anywhere on the lozenge to select the option. Selected options are highlighted.
Presentation (written/recorded/illustrations/options)	On-screen, vertically arranged.
Scoring model	Independent items. Computer-adaptive scoring model with algorithm which selects the subsequent item depending on the response to the preceding item. All items contribute to new estimate of test takers' ability. The updated ability is used to select Part 2 task.
Supplementary information/criteria	• Context and message of each text is clear; text is self-standing. • Context is accessible to the test taker; subject knowledge not required. • All options are realistic and rooted in the text. • The options all paraphrase language in the text. • All options are of a similar length and structure.

16.2.4 Reading Part 1 sample task and screen display

This task is an extract from an *Oxford Test of English Advanced* practice test. To see how the task displays on screen in the test, see the next section.

Read the letter to a local newspaper and choose the correct answer.

To the editor,

In your editorial on the proposed building project in Almond Avenue, local residents are urged to convey their views at a consultation meeting on 30ᵗʰ March, as well as through online questionnaires. Given our experience over the last decade regarding applications for permission to build, the outcome in this case would appear to be a foregone conclusion. The powers-that-be will have decided that a 25-floor apartment block is needed, whatever the implications for traffic, services and amenities in an already overcrowded district. Nonetheless, I will attend this meeting, and all subsequent consultations over the next nine months, and I urge everyone to do likewise. Developers, architects and planning officers need to hear the serious objections we have.

In commenting on the consultation process, the writer expresses …

A concern about the number of people taking part.
B frustration about the timescale involved.
C scepticism about the impact it will have.

Answer key: C

16.2.5 Reading Part 1 screen display

Reading Part 1   01:04

Read the extract from a blog and choose the correct answer.

What was unexpected in the findings from research on learning two languages simultaneously?

- the impact observed on learners' first language ability
- the associated improvements seen in learners' non-language skills
- the extent to which learners reported confusion between languages

www.thelanguagemaster.oup.uk

You'd instinctively presume that learning two languages simultaneously would result in students mixing up one with the other due to their efforts being spread too thinly. A recent investigation suggests otherwise. Students learning a single language were compared with those studying the same language plus another besides. Interference from the mother tongue, a well-documented effect of language acquisition, was evident in both groups to a roughly similar degree. Perhaps surprisingly, those in the latter group were as proficient in both target languages as those in the former were in their single one. Stunning was the fact that double-language students scored more highly relative to the others in logic, reasoning, and mathematical assessments conducted at the end of the research.

Jump to

© Oxford University Press

Next

16.3 Reading Part 2 (Matching) design and sample tasks

Module	Reading
Part	2
Task label	Matching
CEFR	B2–C1

16.3.1 Reading Part 2 task features

	B2	C1
Skill focus	Expeditious search reading at local and global levels: • identify specific information • identify opinion and attitude	Expeditious search reading at local and global levels: • identify specific information • identify opinion and attitude • identify implied meaning
Description	Expeditious search reading to identify specific information, opinion and attitude. This part is timed to ensure that reading is expeditious.	
Number of items per task	Multiple matching: seven questions and three texts OR Matching six profiles of people to match with four longer text descriptions	Multiple matching: seven questions and three texts
Time allocated for task	8 minutes	
Instructions for candidates (task rubric)	Read the questions and match them to the correct text about [insert task-specific text here]. OR The people below are looking for [insert task-specific text here]. Read the information about the people and the [insert task-specific text here] and choose the best [insert task-specific text here] for each person.	Read the questions and match them to the correct text about [insert task-specific text here].
Question format and presentation	Three or four texts presented vertically. Each text comes under one main heading and a subheading. Each text has either a named individual, features a review, or describes a place or location, as a heading. The six or seven questions are presented as direct questions. All options for each question are the same; the options are the names/headings presented above each text.	
Description of the item type (number of words/lexical range – refer to Oxford 5000/grammar/illustrations)	Items = 17 words maximum per question (7 x 17 = 119 maximum total). Lexis in stem and options should avoid C1 or higher lexis as defined within the Oxford 5000. OR Profiles = 45 words maximum (6 x 45 = 270). Lexis in profiles and	Items = 20 words maximum per question (7 x 20 = 140 maximum total). Lexis in stem and options should avoid C2 or higher lexis as defined within the Oxford 5000.

	descriptions should avoid B2 lexis or higher as defined within the Oxford 5000.	
Targeted CEFR level(s)	B2 or C1	
Targeted cognitive functions	☑ Creating intertextual representation	
	☑ Creating a text-level representation	
	☑ Building a mental model	
	☑ Inferencing	
	☑ Establishing propositional meaning	
	☑ Syntactic parsing	
	☑ Lexical access	
	☑ Word recognition	

16.3.2 Reading Part 2 features of the input

	B2	C1
Word count	Three texts, each 190-205 words, plus 119 words for questions. Total word count: 724 words OR Four texts, each 100-125 words, plus 270 words for profiles. Total word count: 770 words	Three texts, each 225-250 words, plus 140 words for questions. Total word count: 890 words
Discourse/text types		
Academic journal extract	☐	☐
Advert	☐	☐
Blog entry	☐	☐
Email	☐	☐
General interest book extract	☑	☑
General interest article extract	☑	☑
Letter	☐	☐
Newspaper editorial	☐	☐
Note	☐	☐
Notice	☐	☐
Professional journal extract	☐	☐
Report extract	☐	☐
Review	☐	☐
Text message	☐	☐
Range of grammatical features present/allowed	Grammatical features appropriate for the level.	No restrictions on range of grammatical features permitted.
Content (specificity, neutrality, concreteness)	Concrete and abstract, including implied meaning.	Concrete and abstract, but with a strong emphasis on implied meaning.
Lexical range (refer to Oxford 5000)	Primarily A1–B2. C1 lexis permitted, although is not tested, as defined in the Oxford 5000.	Primarily A1–C1. C2 lexis permitted, although is not tested, as defined in the Oxford 5000.
Genre/domain (e.g. personal/professional/education/other)	The texts are centred around one topic. The topic is written to be accessible to the target reader. They should be general interest texts on more academic topics. However, no specialized	

	knowledge is required to understand the progression of the texts or the arguments being made within them.

16.3.3 Reading Part 2 features of the expected response

Description	Selection of a single option from three or four available.
Presentation (written/recorded/illustrations/options)	On-screen, vertically arranged.
Scoring model	The six or seven items are grouped. Each item has a difficulty value, and the task has an overall average difficulty value. The algorithm selects tasks based on their average difficulty. The majority of the items are at the same CEFR level as the overall average difficulty (e.g. if task is labelled C1, a minimum of four out of the six items, or five out of seven items will be C1). All items contribute to new estimate of test takers' ability. The updated ability is used to select the Part 3 task. The algorithm will select a task with a different topic to that presented in Parts 1 and 2.
Supplementary information/criteria	Each question covers a topic which is addressed by multiple texts. Therefore, it is necessary to read more than one text for each question to be sure that the answer is correct.

16.3.4 Reading Part 2 sample task – C1 level

This task is an extract from an *Oxford Test of English Advanced* practice test. To see how the task displays on screen in the test, see the next section.

You have 8 minutes to complete this task. Read the questions and match them to the correct book review.

A Brief History of Time

Three reviews of physicist Stephen Hawking's famous book

A Charlene Scott

A definitive work that catapulted its author, the famous physicist Stephen Hawking, into the public consciousness and on to endless bestseller lists, *A Brief History of Time* is often described as the most-bought, least-read book in existence. And that's according to Hawking himself, who, known for his wry sense of humour and ability to poke fun at himself, though these traits do not exactly leap off the page, suggested that many people may have acquired the book simply to boost their intellectual credibility. Hawking set out to make the subject matter of the book – the universe and its structure, origins and evolution – accessible to non-specialists by writing in a non-technical style. An informal survey of friends, several of whom have studied science at a relatively high level, leads me to think the great physicist's success in this respect is somewhat debatable. It is also important to highlight that huge advances have been made in astrophysics since the book was first written in the 1980s, and a number of claims that Hawking then made are now not thought to be correct, or at least are still to be proven. However, among those who are properly equipped to grasp the ins and outs of the 'Big Bang', thermodynamics, black holes, worm holes and so on – such as fellow theoretical physicists – there appears to be almost unanimous admiration for the flair with which Hawking managed to summarize what the universe is comprised of and how it came to be what it is.

B Robert Bradshaw

When Hawking first approached a publisher with his idea for *A Brief History of Time*, he made it clear that he wanted the book to be accessible to everyone and not just a serious work for other physicists. His aim was to establish that, as a result of significant breakthroughs in knowledge through the work of 20th-century physicists and astronomers, humanity was on the verge of finding a unified theory of the cosmos and to 'explain how far we had come in our understanding of the universe'. The editor who read the manuscript told the physicist that general readers would not purchase the publication if he insisted on retaining the many mathematical equations he had included. Hawking agreed to remove all the equations bar one: Einstein's $E = mc^2$. *A Brief History of Time* has now sold more than 25 million copies worldwide. This is far beyond what anyone could have anticipated in 1988, when it was first published. In those days, there was no mass readership for science books generally, let alone one that explored such obscure and complex subjects as black holes, superstrings and the finite yet boundless nature of the universe. Hawking's achievement was to break down the tough theoretical concepts and explain them in plain language, often with the help of deft humour and analogy. So effective was he that many well-known scientists have put their decision to study physics, maths, astronomy or chemistry down to the inspiration they found in Hawking's writing.

C Lucia Marlow

A Brief History of Time is a seminal volume in science writing that explores profound and difficult questions about time and the universe. With great imagination and skill, physicist Stephen Hawking leads us through the mysteries at the heart of creation. He takes us through the history of astronomical studies, paying due recognition to the contributions of astronomers and physicists from the Ancient Greeks such as Aristotle and Ptolemy up to Einstein and fellow giants of the 20th century, and explains in terms within the grasp of non-specialists the reasoning behind some of the most fundamental principles in quantum physics and the theory of relativity. If I had to express one reservation about the book, it would be that, to my mind, not enough of Hawking's own life and personality is evident. And that is to be regretted. Those who knew the man speak of his impish wit, for example, yet I find that missing from the book, though for many that may not be an attractive selling point. That aside, however, his book is a remarkable achievement. In 1988 when it was first published, it was at the cutting edge of what was known about the universe, and recognized as such by many, though not all, of Hawking's peers. Since then, huge advances have been made in astrophysics. However, the book has certainly stood the test of time. Leading physicists highlight its importance, and it is recommended reading for anyone with even a passing interest in astronomy and physics.

1 Who expresses disappointment that Hawking does not reveal much about himself in the book?   **A   B   C**

2 Who mentions that some people bought the book to give a false impression?   **A   B   C**

3 Who refers to scepticism about an early draft of Hawking's book?   **A   B   C**

4 Who approves of the way that Hawking acknowledges the work of other scientists?   **A   B   C**

5 Who mentions the impact the book has had on some people's careers?   **A   B   C**

6 Who questions the extent to which the average reader will understand the ideas discussed in the book?   **A   B   C**

7 Who points out how unlikely it was that a book of this kind would become so popular?   **A   B   C**

Answer key:

1	C
2	A
3	B
4	C
5	B
6	A
7	B

16.3.5 Reading Part 2 screen display

16.4 Reading Part 3 (Gapped text) design and sample tasks

Module	Reading
Part	3
Task label	Gapped text
CEFR	B2 or C1

16.4.1 Reading Part 3 task features

Skill focus	Careful global reading: • identifying text structure • identifying organizational features of a text
Description	Gapped sentences. Identifying text structure and organizational features of a text.
Number of items per task	Six extracted sentences plus one distractor
Time allocated for task	11 minutes
Instructions for candidates (task rubric)	Six sentences are missing from this text. Drag and drop the correct sentence into each gap. There is one extra sentence which you do not need to use.
Question format and presentation	The text is presented on-screen, with the six gaps clearly marked and labelled. Extracted sentences are presented alongside the gapped text. Dragging and dropping sentences in the gaps allows the sentences to be read in context.
Description of the item type (number of words/lexical range – refer to Oxford5000/grammar/illustrations)	Each extracted sentence is 10–30 words in length. Lexis in options may include C1 lexis as defined within the Oxford 5000. Sentences are written to be of approximately equal length.
Targeted CEFR level(s)	B2 or C1
Targeted cognitive functions	☐ Creating intertextual representation ☐ Creating a text-level representation ☑ Building a mental model ☑ Inferencing ☑ Establishing propositional meaning ☑ Syntactic parsing ☑ Lexical access ☑ Word recognition

16.4.2 Reading Part 3 features of the input

	B2	C1
Word count	Text plus extracted sentences (including one option not used): 500–575 words. The text title and glossary are included in the word count.	Text plus extracted sentences (including one option not used): 720–850 words. The text title and glossary are included in the word count.
Discourse/text types		
Academic journal extract	☐	☐
Advert	☐	☐
Blog entry	☐	☐
Email	☐	☐
General interest book extract	☑	☑
General interest article extract	☑	☑
Letter	☐	☐
Newspaper editorial	☐	☐
Note	☐	☐
Notice	☐	☐
Professional journal extract	☐	☐
Report extract	☐	☐
Review	☐	☐
Text message	☐	☐

Range of grammatical features present/allowed	Grammatical features appropriate for the level.	No restrictions on range of grammatical features permitted.
Content (specificity, neutrality, concreteness)	Concrete and abstract	
Lexical range (refer to Oxford 5000)	Primarily A1–B2. C1 lexis permitted, although is not tested, as defined in the Oxford 5000.	Primarily A1–C1. C2 lexis permitted, although is not tested, as defined in the Oxford 5000.
Genre/domain (e.g. personal/professional/education/other)	The text is a long-form article centred around one topic. The topic is written to be accessible to the target reader. It will be a general interest text on a more academic topic. However, no specialized knowledge is required to understand the progression of the text or the arguments being made within it. Technical vocabulary may be included, although if so, it is glossed with an asterisk and a definition of the term offered at the bottom of the article. Individual items do not rely on knowledge of technical terms to answer them correctly.	

16.4.3 Reading Part 3 features of the expected response

Description	Selection of a single option from seven available. Each option can only be selected once.
Presentation (written/recorded/illustrations/options)	Options presented in a list on screen. Options can be dragged and dropped or picked and dropped.
Scoring model	The six items are grouped within a single text. This forms a single Part 3 task. Each item has a difficulty value, and the task has an overall average difficulty value. The selection algorithm selects tasks based on their average difficulty. At least four out of the six items are at the same CEFR level as the overall average difficulty. (E.g. if a task is labelled C1, a minimum of four out of six items will be C1). All items contribute to new estimate of test takers' ability. The updated ability is used to select the next Part 4 item. The algorithm will select a task with a different topic to that presented in Part 3.
Supplementary information/criteria	• The text is written to be coherent and cohesive throughout. • The first paragraph does not include a gap. • Each extracted sentence unambiguously fits only one gap. • Each extracted sentence is attractive as a distractor for at least one other gap. • The extra distractor is plausible and attractive for several gaps. • Each extracted sentence fits with text content both before and after the gap.

16.4.4 Reading Part 3 sample task – C1 level

This task is an extract from an Oxford Test of English Advanced practice test. To see how it displays on screen in the test, see the next section.

Six sentences are missing from this text. Insert the correct sentence (**A–G**) into each gap (**1–6**). There is one extra sentence which you do not need to use. You have 11 minutes to complete this task.

Taste and aesthetics

An introduction to how two philosophers discussed the way we perceive beauty in the world

The term 'aesthetics' came into prominence in the 18th century as a label for the study of artistic experience with German philosopher Alexander Baumgarten. Scottish philosopher David Hume did not use this term but spoke of 'taste', a refined ability to perceive quality in an artwork.

'Taste' might seem completely subjective. _____1 Isn't art just like this? Perhaps you prefer Dickens and Fassbinder, while I prefer Stephen King and Austin Powers; how can you prove that your taste is better than mine? Hume and German philosopher Immanuel Kant both believed that some works of art really are better than others, and that some people have better taste. How could they account for this?

The two took different approaches. Hume emphasized education and experience: people of taste acquire certain abilities that lead to agreement about which artworks are the best. Such people, he felt, will eventually reach a consensus, and in doing so, set a 'standard of taste' which is universal. _____2 He said people of taste must 'preserve minds free from prejudice' but thought no one should enjoy immoral attitudes or 'vicious manners' in art. Sceptics now criticize the narrowness of this view, saying that Hume's judges of taste only acquired their values through cultural programming.

Kant too spoke about judgements of taste but was more concerned with understanding how judgements of beauty are formed. _____3 He tried to describe the human abilities to perceive and categorize the world around us, outlining a complex interplay among mental faculties including perception, imagination, and intellect or judgement. He held that to function in the world to achieve our human purposes, we label much of what we sense, often in fairly unconscious ways. For example, we recognize round flat things and categorize some of these as dinner plates. Then we use them to eat meals. Similarly, we recognize some things as food and others as potential threats.

It is not easy to say how we categorize things like red roses as beautiful. The beauty of the rose is not out there in the world, as the roundness and flatness are in the plate. _____4 And yet there is some basis for claiming that the roses are beautiful. After all, there is quite a lot of human agreement that roses are beautiful and cockroaches are ugly. Hume tried to resolve this problem by saying that judgements of taste are 'intersubjective': people with taste tend to agree with each other. Kant believed that judgements of beauty were universal and grounded in the real world, even though they were not actually objective. How could this be?

Kant noted that we typically apply labels or concepts to the world to classify sensory inputs that suit a purpose. _____5 Beautiful objects, however, do not serve ordinary human purposes, like feeding ourselves. A beautiful rose pleases us, but not because we want to eat it. His way of recognizing this was to say that something beautiful has 'purposiveness without a purpose'. This curious phrase needs further unpacking.

When I perceive a rose as beautiful, this is not quite like putting it into my mental cupboard of items labelled 'beauty' – nor do I just throw the cockroach into my mental trash can of 'ugly' items. But features of the object almost force me to label it as I do. The rose might have its own purpose, to reproduce new roses, but that is not why it is beautiful. _____6 This is what Kant means by saying that beautiful objects have purpose. We label an object beautiful because it promotes an internal harmony or 'free play' of our mental faculties; we call something 'beautiful' when it elicits this pleasure. When you call a thing beautiful, you thereby assert that everyone ought to agree.

Adapted and abridged from an authentic source, acknowledged at the end of the test.

A These experts, in his view, can differentiate objects of high quality from less good creations.

B If it did have such exact attributes, we would surely not get into so many disagreements regarding taste.

C The aesthetic, therefore, is experienced when a sensuous object stimulates our emotions, intellect and imagination.

D Some people have favourite colours and desserts, just as they favour certain kinds of automobiles or furniture.

E Something about that particular combination of colour and texture prompts my mental faculties to feel that what I'm looking at is 'right.'

F Good ones, it was felt, are grounded in features of artworks and objects themselves, not just in ourselves and our preferences.

G For example, on finding a cylindrical drinking vessel in the dishwasher that I recognize as a mug, I put it in the cupboard with other mugs, not with spoons.

Answer key:

1	D

2	A
3	F
4	B
5	G
6	E

16.4.5 Reading Part 3 screen display

Reading Part 3   10:54

Six sentences are missing from this text. Drag and drop the correct sentence into each gap. There is one extra sentence which you do not need to use.

We want to feel that our world is comprehensible so we can be responsible for it, and we feel alienated by impersonal, strange forces.

Yet designing, making and selling new technical products involves numerous people – many of whom are not engineers.

Nevertheless, many people still regard the products of engineering as morally and ethically neutral – they are inherently neither bad nor good.

Heat sources and electricity, not to mention convenient ways to travel, are all aspects of life that we take for granted until one of them lets us down at any given moment.

Transport is a good example: the canal network opened up possibilities for trade, while the steam engine and the railways created new opportunities for travel.

Since the Industrial Revolution in Britain in the 18th century, however, the interior workings of many have become mysterious, complex and non-transparent to all but the experts.

Just as a healthy lifestyle is easier if you have empathy with

Engineering

Engineering is, in its most general sense, turning an idea into a reality – creating and using tools to accomplish a task or fulfil a purpose. The ability of humans to make tools is remarkable. But it is our ingenious ability to make sense of the world and use our tools to make even more sense and even more ingenious tools that makes us exceptional. Tools are part of what it is to be human.

However, for many, there is a disturbing cloud. Once, tools were simple common sense – almost all were understandable to the non-specialist. _____ In fact, engineers have been accused of 'rendering many of the devices we depend on every day unintelligible to direct inspection', and there is a realization that science, technology, engineering and mathematics are interwoven in a way that needs separating. This is urgent because we need to understand what engineering offers, what it might offer in the future, and what it cannot ever deliver.

All through history, people have expressed their awe, spirituality and faith by making tools. From ancient stone tablets to present-day Liquid Crystal Display (LCD) screens, from horse-drawn carts to space rockets, our tools have become ever more powerful, conspicuous and central to our lives. Imagine waking up in the morning, attempting to switch on your light, but finding that everything provided by engineers had gone. Many disaster movies rely on this kind of idea. _____ At the beginning of the 21st century, almost everything we use is a product of the activities of engineers and scientists.

Of course, just because you rely on something doesn't mean you

⚙ Jump to   © Oxford University Press   Next

16.5 Reading Part 4 (Long text) design and sample tasks

Module	Reading
Part	4
Task label	Long text
CEFR	B2 or C1

16.5.1 Reading Part 4 task features

	B2	C1
Skill focus	Careful local and global reading: • identifying facts/attitude/opinion • identifying writer's purpose • understanding words in context • understanding reference words • understanding global meaning • understanding rhetorical purpose (exemplification)	Careful local and global reading: • identifying facts/attitude/opinion • identifying writer's purpose • understanding implication • understanding rhetorical purpose (exemplification, comparison, exaggeration)
Description	Multiple-choice questions. This task focuses on propositional and structural meaning, such as identifying attitude/opinion, purpose, understanding implication, exemplification, reading for gist, inference, and global meaning.	
Number of items per task	Four or five three-option multiple-choice questions on a single text.	
Time allocated for task	8 minutes	
Instructions for candidates (task rubric)	Read the passage about [insert task-specific text]. For Questions 1–4 choose the correct answer OR Read the article about [insert task-specific text]. For Questions 1–5, choose the correct answer.	
Question format and presentation	The text is presented on-screen, with the four or five questions presented alongside. The questions are all visible simultaneously.	
Description of the item type (number of words/lexical range – refer to Oxford 5000/grammar/illustrations)	Four items (stem and three options) = maximum 200 words. Lexis in stem and options should avoid C2 lexis as defined within the Oxford 5000.	
Targeted CEFR level(s)	B2 or C1	
Targeted cognitive functions	☐ Creating intertextual representation	
	☑ Creating a text-level representation	
	☑ Building a mental model	
	☑ Inferencing	
	☑ Establishing propositional meaning	
	☑ Syntactic parsing	
	☑ Lexical access	
	☑ Word recognition	

16.5.2 Reading Part 4 features of the input

Word count	Text: 350 words. The text title and subtitle are included in the word count.	Text: 630–680 words. The text title, subtitle and glossary are included in the word count.
Discourse/text types		
Academic journal extract	☐	☐
Advert	☐	☐
Blog entry	☐	☐
Email	☐	☐
General interest book extract	☑	☑
General interest article extract	☑	☑
Letter	☐	☐
Newspaper editorial	☐	☐
Note	☐	☐
Notice	☐	☐
Professional journal extract	☐	☐
Report extract	☐	☐
Review	☐	☐
Text message	☐	☐
Range of grammatical features present/allowed	Grammatical features appropriate for the level.	No restrictions on range of grammatical features permitted.
Content (specificity, neutrality, concreteness)	Concrete and abstract, including implied meaning.	Concrete and abstract, but with a strong emphasis on implied meaning.
Lexical range (refer to Oxford 5000)	Primarily A1–B2. C1 lexis permitted, although is not tested, as defined in the Oxford 5000.	Primarily A1–C1. C2 lexis permitted, although is not tested, as defined in the Oxford 5000.
Genre/domain (e.g. personal/professional/education/other)	General interest	The topic is of general academic interest to test takers, but does not require specialist knowledge to comprehend. Technical vocabulary may be included, although if so, it is glossed with an asterisk and a definition of the term offered at the bottom of the article. Individual items do not rely on knowledge of technical terms to answer them correctly.

16.5.3 Reading Part 4 features of the expected response

Description	Selection of a single option from three available.
Presentation (written/recorded/illustrations/options)	On-screen, vertically arranged.
Scoring model	The four or five questions are grouped around a single text. Each item has a difficulty value, and the task has an overall average difficulty value. The selection algorithm selects tasks based on their average difficulty. All items will be at the same CEFR level as the overall average difficulty. All items contribute to the final

	estimate of test takers' ability, which is then reported as the test taker's reading ability.
Supplementary information/criteria	- The text is written to be coherent and cohesive throughout. - The test taker needs to read the text to answer the task (testing language skills, not world/subject knowledge). - The options are of a similar length and structure. - The options all paraphrase language in the text, where possible. - All options are realistic and rooted in the text. - Questions are presented in the same order as the information in the text.

16.5.4 Reading Part 4 sample task – C1 level

This task is an extract from an *Oxford Test of English Advanced* practice test. To see how the task displays on screen in the test, see the next section.

Read the passage about cinema. For questions 1–5, choose the correct answer. You have 8 minutes to complete this task.

The end of cinema?

The idea that cinema is dying is nothing new; the death of cinema has been announced many times before, starting in the 1920s when sound was first added to film. There has been a real flood of such claims of late, however, and for an unprecedented number of reasons. In comparable cases, such as the predicted death of the novel, announcements of this kind seem to have led to nothing but revival after revival. Nonetheless, something is happening to provoke these assertions and arguments, and it's worth trying to find out what it is.

For some film lovers, the end began in the 1990s with what looked like the disappearance of a film type they could admire: movies comprising a curious mixture of commerce and art that had been plentiful since the 1950s. For other people, the end was not about film content or style but about technology, and specifically the possibility of seeing films at home. But hadn't television embodied the same issue with its introduction into domestic life in the 1950s? Not quite, although it had provided serious competition and had somewhat reduced audiences and driven the movie sector to do things with large screens and computer-generated imagery that only cinemas could properly display.

Television by itself didn't alter things that much, however. There were still fixed schedules on television and fixed schedules in commercial theatres – no one was interfering with film time. It was the invention of video recorders, and mass ownership of them in the 1980s, that brought the most radical changes. People could now skip whole sequences without having to leave the room. They no longer needed to watch shows when they were screened; theoretically, programming became irrelevant. And, by buying or renting videos, people could watch films exclusively in their own living rooms.

Developments in digital technology since the 1990s have had a massive impact on film and cinema. Nevertheless, cinema remains big business. The highest grossing films still make multi-million-dollar profits and most films we see in the cinema continue to be films in the old sense. Films are also cheaper to make, and more are being produced than ever. The best way to see a great film may still be on the big screen, but what is currently available for free on any smart phone probably has better definition than any version available through more traditional media such as VHS or analogue TV. And the more dominant digital technology becomes, the more complex the situation is. Sometime today, someone you know will film some moving images, and upload them to social media. This will be a piece of film but won't be on film. And it won't be seen in a cinema.

Whatever the long-term future of the cinema as an institution may be, I believe the notion of film will survive in two senses. There is the dictionary definition of a film as 'a cinematographic representation of a story, drama, episode, event, etc.' When people say they have been watching a film, this is what we understand they mean – and we probably always will – and a film-maker is someone who makes such things. The other sense is that of fragments or sequences, short or long, fictional or actual, of motion caught in the act. These may or may not add up to a film in the traditional sense, but it's part of the social-media-driven freedom of the form that they don't have to.

Why should we care about the future of cinema? Perhaps it's because film appears to fit so perfectly with what scholars Leo Charney and Vanessa Schwartz describe as our 'impulse to define, fix and represent isolated moments in the face of modernity's distractions and sensations'. But surely it does that only when we treat it as a form of rolling photography? Film at its magical, animating best doesn't define, fix, capture or even necessarily represent anything. Instead, it lets its subjects run, gives them back the life we thought they had lost; and mixes what we remember with what we dream.

Adapted and abridged from an authentic source, acknowledged at the end of the test.

1 In the first paragraph, the writer suggests that recent threats to cinema …

A prompt a sense of hope for the future of the art form.

B are of greater significance than any previously encountered.

C have their roots in developments that originated about a century ago.

2 The writer refers to special effects in order to …

A indicate the commercial pressures present in the industry.

B illustrate the ability to incorporate technical innovations.

C highlight the shift away from a focus on artistic quality.

3 In the third paragraph, the writer suggests that, before the 1980s, television …

   A had a less dramatic impact than had initially been hoped for.
   B represented part of a wider trend towards home-based entertainment.
   C lacked certain features that would encourage people to stay away from cinemas.

4 In comparing big screens and smart phones, the writer makes the point that …

   A the ability to view new content digitally is negatively impacting cinema.
   B the type of content that people are choosing to watch is evolving.
   C the influence conventional cinema has over new content is profound.

5 The writer thinks that Charney and Schwartz's words …

   A reflect an overly analytical approach to thinking about film.
   B signify an outdated view of the potential that film offers.
   C reveal a limited understanding of the appeal film can have.

Answer key:

1	B
2	A
3	C
4	B
5	C

16.5.5 Reading Part 4 screen display

17 Appendix 4: Writing module details

17.1 Writing module construct and domain definitions

Test purpose	The *Oxford Test of English Advanced* Writing test assesses test takers' written English proficiency for two sub-domains: academic and professional purposes. Within these sub-domains, the test assesses the ability to: • write clear, well-structured expositions which develop an argument; offer reasons in support of or against a particular point; and evaluate different ideas or solutions to a problem • identify main ideas; synthesize, summarize and structure them into a coherent and concise text; paraphrase and redraft text materials typically encountered in university settings.	
Examples of domain activities	Writing for academic and professional purposes: relevant text genres include reports, abstracts/summaries (for classmates), essays (for tutors), note taking. Reading into writing – identifying the main ideas in more than one text on the same topic and synthesizing the information for use by others, either for a report or for an essay.	
	Essay (academic)	**Summary (professional and academic)**

Communication goals by sub-domain	• Organizing text to develop ideas and arguments	• Summarizing and synthesizing information from two sources
	• Expressing opinion with clarity and precision	• Identifying relevant information
	• Structuring an argument with suitable introduction and conclusion	• Distinguishing main from supporting information
	• Giving reasons in support of or against a particular point	• Structuring a text
	• Evaluating different ideas or solutions to a problem	• Paraphrasing and circumlocution
Targeted cognitive functions	**Encoding and execution:** the test taker transforms ideas into words, clauses, sentences, paragraphs and text using linguistic resources such as vocabulary (lexicalization) and grammar (grammatical encoding).	
	Discourse organization: conceptualizing, connecting and developing ideas using organizational devices to cohesively and coherently structure a written text and show awareness of topic, goals and target audience.	
	Monitoring and revising: the test taker re-reads what has been written and revises accordingly, e.g. new ideas may be added, structure amended, overall message adjusted, wording changed to express an idea better, spelling or grammar or keystroke errors corrected, circumlocution and/or paraphrasing used to make text more original.	

17.2 Writing Part 1 (Essay) design and sample tasks

Module	Writing
Part	1
Task label	Essay
CEFR	B2–C1

17.2.1 Writing Part 1 task features

Skill focus	• Organizing text to develop ideas and arguments • Writing clear, well-structured expositions • Expressing opinion with clarity and precision
Description	Test taker writes a response to written prompt. Test takers are presented with an essay question and three ideas, of which at least two must be included in the essay. Test takers are required to write an essay for an audience (their tutor), structuring with an introduction, a set of well-expanded points in support of an argument, and a conclusion.
Number of items per task	One
Time allocated for task	30 minutes
Instructions for candidates (task rubric)	**The clock shows how much time you have to write an essay. Write 220–280 words.** You have been discussing the topic of […] in your class. Your tutor has asked you to write an essay on the following: [Essay statement] [Question]

	Your essay must include at least **two** of the following ideas: [Three bulleted prompts] Organize your essay clearly, introducing the topic, providing support for the points you make, and giving a conclusion. Write your **essay**.
Question format and presentation	The topic will always be directly related to the essay statement, or may be a verbatim extract. The essay statement consists of a single sentence or a pair of opposing sentences. It is always followed by a question, which has been designed to ensure the test taker has the opportunity to argue a case. The essay statement and question are presented in a box.
Targeted CEFR level(s)	B2–C1
Targeted cognitive functions	☑ Revising
	☑ Monitoring
	☑ Execution
	☑ Encoding
	☑ Planning: micro and macro
	☑ Conceptualization

17.2.2 Writing Part 1 features of the input

Description	The essay statement will be a straightforward sentence or pair of opposing sentences sufficiently wide in scope to support the three bulleted prompts which follow, allowing the test taker to develop an argument and give a conclusion. The essay statement is always followed by one of the questions below: • Do you agree or disagree? • How far do you agree or disagree with this statement? • Do you think that this is a positive or a negative [situation/development (etc.)]? • Do the advantages of this [situation/development (etc.)] outweigh the disadvantages? • Is this a good or a bad thing? • Which opinion do you agree with? Three **bulleted prompts** then follow.
Word count	80 words (maximum) (inclusive of essay statement, question and three bulleted prompts, exclusive of task rubric) Bulleted prompts: 6 words each (maximum)

Grammatical features	• The task rubric is standard.
	• Bulleted prompts are short and simple. They must be:
	  ○ Noun phrases providing opportunities for the development and exemplification of points
	  ○ Broad in scope; not directed towards answering the essay in a particular way

Content (specificity, neutrality, concreteness)	Concrete/abstract
Lexical range (refer to Oxford 5000 for CEFR levels of individual words)	Task will be composed of A1–B2 lexis so that it is accessible to B2-level test takers. Task has sufficient scope for C1 candidates to show their language ability.

Genre/domain (e.g. personal/professional/education/other)	Academic/professional

17.2.3 Writing Part 1 features of the expected response

Description	220–280 words. Test takers produce a clearly structured essay, arguing a case with an introduction and conclusion.
Presentation (written/recorded/illustrations/options)	Essay task, essay statement, question and bulleted prompts appear on the left-hand side of the screen. A text box is on the right-hand side for the test taker to produce their response. There is a word counter allowing the test taker to see how many words they have written. There is a timer, counting down from 30 minutes, indicating how much time remains. Upon reaching zero, the test automatically progresses to Part 2, saving the response.
Discourse functions	☑ Organizing text to develop ideas and arguments
	☑ Expressing opinion with clarity and precision
	☑ Structuring an argument with suitable introduction and conclusion
	☑ Giving reasons in support of or against a particular point
	☑ Evaluating different ideas or solutions to a problem
Scoring model	Part 1 is marked independently of Part 2. Test takers receive four marks (for Task fulfilment, Organization, Grammar, and Lexis) for each script. The marks from Script 1 (Part 1) and Script 2 (Part 2) are added together. The total is then translated into a standardized score and a CEFR level.
Supplementary information/criteria	• Statements will not contain multiple clauses or be ambiguous.
	• When an essay statement consists of a pair of opposing sentences, the following frame is used:
	o Some people say (that) […]. However, others argue that […].
	o Which opinion do you agree with?

17.2.4 Writing Part 1 sample task

This task is an extract from an *Oxford Test of English Advanced* practice test. To see how the task displays on screen in the test, see the next section.

You have 30 minutes to write an essay. Write 220–280 words.

You have been discussing the topic of fashion in your class. Your tutor has asked you to write an essay on the following:

> *Some people say that fashion stores should not sell cheap clothing which does not last long. However, others argue that it is a good thing that everyone can afford to follow the latest fashions.*
>
> *Which opinion do you agree with?*

Your essay must include at least **two** of the following ideas:

- impact on businesses
- impact on consumers
- impact on the environment.

Organize your essay clearly, introducing the topic, providing support for the points you make, and giving a conclusion.

Write your **essay**.

17.2.5 Writing Part 1 screen display

17.3 Writing Part 2 (Summary) design and sample tasks

Module	Writing
Part	2
Task label	Summary
CEFR	B2–C1

17.3.1 Writing Part 2 task features

Skill focus	• Expeditious and careful local reading • Synthesizing the main ideas contained in two texts of different text types/genres into a well-structured summary
Description	Test taker writes a summary of two input texts. This is an integrated-skills, intertextual reading-into-writing task.
Number of items per task	One
Time allocated for task	20 minutes
Instructions for candidates (task rubric)	**The clock shows how much time you have to write a summary. Write 80-100 words.**

| | [Task rubric] You have been learning about an aspect of [general topic] for a college course. You have read a textbook extract and attended a lecture and now your tutor has asked you to write a summary of the main ideas for your classmates to read.

[Standard rubric] Write **one** paragraph, combining information from the textbook extract and the lecture transcript to summarize the **main ideas**. Your summary should provide the reader with enough information to understand the main ideas from both texts.

Write full sentences, using **your own words** where possible. Do **NOT** write more than 100 words.

Write your **summary**.

[Textbook extract] [Lecture transcript]

[Glossary] |
|---|---|
| **Question format and presentation** | For each task the rubric is divided into two. There is a task rubric followed by the standard rubric. The standard rubric is always the same. The task rubric will always differ for each summary. After the task and standard rubrics, the two input texts appear alongside each other. A glossary appears beneath the texts. Guidance on the content of the texts is outlined in 'features of the input'. |
| **Targeted CEFR level(s)** | B2–C1 |
| **Targeted cognitive functions** | ☑ Revising |
| | ☑ Monitoring |
| | ☑ Execution |
| | ☑ Encoding |
| | ☑ Planning: micro and macro |
| | ☑ Conceptualization |

17.3.2 Writing Part 2 features of the input

Description	- Two texts on the same subject. The subject will be academic or professional. - Texts are propositionally complex but remain accessible to the general reader. One text is a textbook extract and one text is a transcript of an extract from a lecture. Each text genre is clearly marked in the text heading. - There is an overarching main idea which unites both texts. - There will be a further two or three main ideas spread across the two texts. Each of these main ideas has at least one illustrating example or other supporting detail. - The texts are designed to encourage test takers to synthesize information from both when summarizing; for example, a supporting detail relating to a main idea may appear in a different text to the main idea itself. - The number of supporting details for each of the main ideas is not fixed. - There will be some redundancy in the texts (i.e. information that it is not necessary to include in the summary).

	• Each text is presented in a single paragraph.
	• There will be a glossary of less frequent lexis (i.e. B2 plus and above) (3–6 glossed words). Words targeted for glossing will be those which are required to understand the text but will not be key topic words required for the summary. Definitions will use B2 lexis and below.
Word count	Input: 130–170 words per text. 300 words total (maximum) 280 words (minimum) Word counts do not include text headings.
Grammatical features	• The texts should appear to be extracts from longer texts, meaning that discourse features may indicate continuation from an earlier text (not present in the task). • The Spoken text will contain some instances of features which clearly mark it as spoken text (for example false starts, repetition, ellipsis, self-correction). • Grammatical structures need to be accessible to B2-level test takers.
Content (specificity, neutrality, concreteness)	Concrete/abstract
Lexical range (refer to Oxford 5000 for CEFR levels of individual words)	Input texts will be composed primarily of A1–B2 lexis so that they are accessible to B2-level test takers. The academic/professional subject matter means that some higher-level lexis will be included. This lexis will be glossed either within the texts themselves or in the glossary.
Genre/domain (e.g. personal/professional/education/other)	Academic

17.3.3 Writing Part 2 features of the expected response

Description	80–100 words (maximum). Test taker summarizes the main ideas from two texts of different text types/genres by synthesizing information into a well-structured summary. Redundancy requires the test taker to demonstrate their ability to identify relevant information and distinguish main from supporting information, which is part of the construct of the task.
Presentation (written/recorded/illustrations/options)	The Summary task rubric and standard rubric appear on the left-hand side of the screen, the input texts at the bottom of the screen, and a text box is on the right-hand side for the test taker to write their response. There is a word counter allowing the test taker to see how many words they have written. The glossary can be accessed using the 'Glossary' button in the ribbon at the bottom of the screen. There is a timer, counting down from 20 minutes, indicating how much time remains. Upon reaching zero, the test automatically progresses to the end of the test, saving the response.
Discourse functions	☑ Summarizing and synthesizing information from two sources
	☑ Identifying relevant information
	☑ Distinguishing main from supporting information
	☑ Structuring a text
	☑ Paraphrasing and circumlocution

Scoring model	Part 2 is marked independently of Part 1. Test takers receive four marks (for Task fulfilment, Organization, Grammar, and Lexis) for each script. The marks from Script 1 (Part 1) and Script 2 (Part 2) are added together. The total score is then translated into a whole CEFR level. This is reported as the test taker's writing ability.
Supplementary information/criteria	• The overarching main idea will not always form part of a topic sentence for either text (i.e. appear first in either text). • The location of the main ideas and supporting details within the input texts will vary. • For each summary, assessors can refer to the task-specific marking guide which shows the main ideas and supporting details for each text. This allows assessors to quickly identify which ideas are represented in the test taker's response and the extent to which they have synthesized information from both texts.

17.3.4 Writing Part 2 sample task

This task is an extract from an *Oxford Test of English Advanced* practice test. To see how the task displays on screen in the test, see the next section.

You have 20 minutes to write a summary. Write 80–100 words.

You have been learning about an aspect of the weather for a college course. You have read a textbook extract and attended a lecture and now your tutor has asked you to write a summary of the main ideas for your classmates to read.

Write **one** paragraph, combining information from the textbook extract and the lecture transcript to summarize the **main ideas**. Your summary should provide the reader with enough information to understand the main ideas from both texts.

Write full sentences, using **your own words** where possible. Do **NOT** write more than 100 words.

Textbook extract

In our earlier case study showing how Storm Desmond affected the Lake District in December 2015, the record rainfall was caused by a slow-moving weather front passing over northern England. Weather fronts are the forces behind spells of bad weather in the UK, and they can be divided into three main types. When a warm air mass meets and rises over a cooler mass of air, this is called a warm front. Warm fronts are often followed by cold fronts, in which cold air replaces and pushes up warmer air. At some point, the cold front is likely to catch up with the warm front and this is known as an occluded front. Although the effects of weather fronts are rarely as severe as those seen in the Lake District in 2015, such extreme weather is occurring more frequently than it used to as a direct result of climate change.

Lecture transcript

'Have any of you ever wondered why British weather can change so much from one day to the next? Or even hour by hour? Well, we can blame the constantly changing nature of our weather on areas of high and low pressure, and by pressure I mean the force of the atmosphere on the earth's surface. Areas of low pressure are called depressions, and they bring with them the weather fronts which we are so familiar with in the form of dark clouds and sudden changes in wind and temperature. Of course, the weather isn't always poor here. We often have periods of hot, sunny weather in the summer, or cold days in the winter with clear bright skies – these much calmer conditions appear when an anti-cyclone, or area of high pressure, is nearby.'

Glossary

case study something you read about as an example of an idea or theory you are learning about

mass a large amount of something

spell a period when a particular type of weather lasts

Write your **summary.**

17.3.4.1 Task-specific marking guide

When marking Task fulfilment, please refer to this table.

	Main idea	Supporting details
1	'Areas of high and low pressure' cause changing weather in Britain **(2)**	N/A
2	Depressions (areas of low pressure) are associated with weather fronts which bring bad weather **(2)**	Warm front: 'when a warm air mass meets and rises over a cooler mass of air' **(1)** Cold front (often follows a warm front): 'cold air replaces and pushes up warmer air' **(1)** Occluded front: when a cold front catches up with a warm front **(1)**
3	Anti-cyclones (areas of high pressure) bring calmer conditions **(2)**	Summer: hot and sunny **(2)** Winter: cold with clear bright skies **(2)**

Key
(1) From the textbook extract
(2) From the lecture transcript

17.3.5　Writing Part 2 screen display

Writing Part 2

19:56

The clock shows how much time you have to write a summary.
Write 80–100 words.
You have been learning about an aspect of science for a college course. You have read a textbook extract and attended a lecture and now your tutor has asked you to write a summary of the main ideas for your classmates to read.
Write **one** paragraph, combining information from the textbook extract and the lecture transcript to summarize the **main ideas**. Your summary should provide the reader with enough information to understand the main ideas from both texts.
Write full sentences, using **your own words** where possible.
Do **NOT** write more than 100 words.
Write your **summary**.

Type your answer here …

0 words

Textbook extract

The law of Conservation of Energy states that energy cannot be created or destroyed. It can only be transferred to a different place or changed into a different type of energy. This law explains why some chemical processes produce heat and light. Since the total amount of energy cannot change, the excess energy from these exothermic reactions is released as heat or light. Corrosion – when a metal breaks down due to its environment – is an example of this, but we don't notice the heat as it is released very slowly. Changes of state can also be exothermic. Most people don't realize that when water freezes it is giving off heat, but, of course, heat lost from the water as it cools must be transferred to the surrounding air. The law of Conservation of Energy also explains why some processes take in heat and light. For example, during photosynthesis, plants absorb energy in the form of sunlight and convert it into food allowing

Lecture transcript

'The experiment in front of me is not … er … particularly complicated but it shows you the way that … shows you different types of energy transfer. The gas burning in a flame here is an example of combustion. It is giving off heat and light so this is an exothermic reaction. You can all see the energy released as light and we are about to see the effect of the heat energy on this water above it. If you look carefully, you can see the water is changing into steam which, of course, we call evaporation. The water is absorbing the heat energy, so this is an endothermic process with energy being taken from the environment. Changes of state from solid to liquid to gas require heat so they are always endothermic.'

Jump to　Glossary

Next

18 Appendix 5 – Oxford Test of English Advanced Speaking Script 1 and Script 2 marking criteria

CEFR		Marks	Task fulfilment	Pronunciation and fluency	Grammar	Lexis
C2	C2.1	7	• task requirements **completely** fulfilled • well-developed and clearly communicated message which appears **effortless** • register is **effortlessly** appropriate to purpose of task, audience and context • (Part 3 summary) concisely **synthesizes** the two main points; supporting details (if included) **organized effectively**	• can **exploit** stress, rhythm and intonation effectively to convey meaning **effortlessly** • can express themselves **at length** with a natural **unhesitating** flow • can use a **wide variety** of **cohesive features** to link ideas efficiently and produce **highly organized** discourse	• maintains **consistent** grammatical **control** of even the most complex language forms; errors are **rare** and/or **only** concern complex forms • mastery of **a very wide range** of structures; no sign of restricting their use of structures	• can use a **broad** and **sophisticated** lexical repertoire always relevant to the tasks; errors are **rare** and/or **only** concern infrequent lexis • **consistently appropriate** use of infrequent collocations, idioms, colloquialisms and/or **awareness** of **connotative** levels of meaning
C1	C1.2	6	Meets the positive descriptors in C1.1 comfortably, and the negative descriptors in C1.1 minimally if at all			
	C1.1	5	• task requirements **mostly** fulfilled • listener is **fully** informed within the context of the task • register is **appropriate** to purpose of task and audience; lapses are rare • (Part 3 summary) **synthesizes** the two main points; supporting details (if included) are used **appropriately**	• can **control** stress, rhythm and intonation to convey meaning; lapses do not affect meaning • can express themselves **spontaneously** with a **smooth** flow • can use a **variety** of **cohesive features** to produce **organized** discourse	• maintains a **high level** of grammatical **control** of both simple and complex structures; errors are **rare** and/or **difficult to spot** • a **wide range** of structures to express viewpoints clearly and fully	• uses a **broad** lexical repertoire relevant to the tasks; errors are **rare** and/or **difficult to spot** • can use a **range** of collocations, idioms and/or colloquialisms; **little evidence** of searching and avoidance strategies
B2	B2.2	4	Meets the positive descriptors in B2.1 comfortably, and the negative descriptors in B2.1 minimally if at all			
	B2.1	3	• task requirements **generally** fulfilled • listener is **adequately** informed within the context of the task; points made are not always sufficiently expanded • register is **generally** appropriate to purpose of task and audience; lapses occur • (Part 3 summary) **paraphrases** the two main points; supporting details (if included) may be used **inappropriately**	• can **employ** stress, rhythm and intonation to **support** the meaning with some influence from other language(s); lapses may affect meaning • can express themselves in **stretches** of language with an **even tempo** • can use a **limited variety** of **cohesive features** to link ideas into **coherent** discourse	• maintains **control** of **adequate range** of structures; complex structures used **inaccurately** or **rigidly**; **noticeable errors** occur which may impede understanding • can use **some complex** structures to express viewpoints	• **adequate** range of lexis to express viewpoints on everyday topics; **noticeable errors** occur and **may cause some confusion** • can use some **appropriate** collocations and/or expressions, lexical gaps cause **hesitation** and circumlocution
B1	B1.2	2	Meets the positive descriptors in B1.1 comfortably, and the negative descriptors in B1.1 minimally if at all			
	B1.1	1	• task requirements **partly** fulfilled • listener is only **partially** informed within the context of the task; ideas are not often expanded; tasks may be misunderstood or left unresolved • **limited awareness** of register for task and audience • (Part 3 summary) presents **some relevant** content (main or supporting points)	• can **convey** message, but stress, rhythm and intonation are **influenced** by other language(s); lapses affect meaning • can express themselves in **short contributions**, though **pausing** and repair **very evident** • can use a **limited number** of cohesive features to connect a **sequence** of points	• maintains **control** of **simple** grammatical structures; **noticeable errors** occur which **sometimes** impede understanding • **limited range** of structures; a repertoire of **frequent 'routines'**	• **sufficient** lexis for routine responses to more familiar tasks; errors may be **frequent** for **unfamiliar topics** • **limited repertoire** of **frequent patterns** used repetitively, some **hesitation** even on everyday topics
N/A		0	Response does not meet all the positive descriptors in B1.1 **OR** response meets the caps/limitations outlined below			

	Caps for irrelevant/non-responses (across all criteria)			Caps for Task fulfilment	
	No penalty for irrelevant/non-responses to:	Reduce by **one mark** across all four criteria for irrelevant/non-responses to:	Give **0 marks** across all four criteria for irrelevant/non-responses to:	Cap at **B2.2** for Task fulfilment:	
Script 1	• up to **one** of the Part 1 Interview questions	• **two or three** of the Part 1 Interview questions	• **all** of Part 1 (Interview) • **OR all** of Part 2 (Voice message) • **OR all** of Part 3 (Summary)	**Script 1**	N/A
Script 2	• up to **one** of the Part 5 Follow-up questions	• **two or three** of the Part 5 Follow-up questions	• **all** of Part 4 (Debate) • **OR all** of Part 5 (Follow-up questions)	**Script 2**	• Using **fewer than two ideas** in the Part 4 (Debate)

19 Appendix 6 – Oxford Test of English Advanced Writing Script 1 (Essay) marking criteria

CEFR		Marks	Task fulfilment	Organization	Grammar	Lexis
C2	C2.1	7	- essay is **well developed**; arguments are sophisticated and response appears **effortless** - register is **consistently effective** to purpose of task, audience and context - overall impact on the reader is **completely positive**, conveying argumentative nuances precisely	- ideas are **well organized** into an **effective** and **logical** structure - the text is **highly coherent**, using a **variety of organizational features** - uses a **wide range** of **cohesive features** to produce clear language with a **natural smooth flow**	- mastery of a **wide range** of simple and complex structures - maintains **consistent** grammatical control of even the most complex language forms - maintains a **high level of accuracy throughout**; errors are **rare** and **almost always** concern complex forms	- uses a **wide range of sophisticated** lexis and phrases - uses lexis and phrases **always appropriate** to task purpose with a **high level of accuracy** - errors are **rare** and/or **almost always** concern infrequent lexis or phrases
C1	C1.2	6	Meets the positive descriptors in C1.1 comfortably, and the negative descriptors in C1.1 minimally if at all			
	C1.1	5	- essay is **well developed**; expands and supports points at length - register is **nearly always appropriate** to purpose of task and audience - overall impact on the reader is **consistently positive**	- ideas are **well organized** - the text is **consistently coherent** and well structured - uses a **range** of **cohesive features appropriately** with rare instances of misuse	- a **wide range** of structures to express viewpoints clearly and fully - maintains a **high level** of grammatical control of simple and complex structures - maintains a **good level** of grammatical accuracy throughout; errors are **rare** and **difficult to spot**	- uses a **wide range** of lexis and phrases - uses lexis and phrases **appropriate** to task purpose and with a **good level of accuracy** - any errors are **rare** and usually non-impeding
B2	B2.2	4	Meets the positive descriptors in B2.1 comfortably, and the negative descriptors in B2.1 minimally if at all			
	B2.1	3	- essay is **generally well developed**; develops an argument and points are reasonably **expanded** - register is **generally appropriate** to purpose of task and audience **majority of the response** is understood with **minimal effort** on the part of the reader	- organization of ideas is **generally good** - the text is **generally coherent**, with only occasional breaks in communication - uses a **reasonable range** of appropriate **cohesive features** with only **occasional misuse or overuse**	- has a **good range** and command of simple language structures - uses **some complex** grammatical structures, although these may be used **unnaturally or inaccurately** - grammatical errors **occasionally** impede communication	- uses a **good range** of lexis and phrases - uses lexis and phrases generally **appropriate** to task purpose, although these may be used **unnaturally or inaccurately** - lexical errors **occasionally** impede communication
B1	B1.2	2	Meets the positive descriptors in B1.1 comfortably, and the negative descriptors in B1.1 minimally if at all			
	B1.1	1	- essay is **under developed**; points made are **short** and minimally expanded - register is **sometimes appropriate** to purpose of task and audience - the response is **understood with effort** required on the part of the reader	- organization of ideas is **adequate** - the text is **minimally coherent**, with breaks in communication common - uses **simple cohesive features**, such as 'and', 'but' and 'because'	- evidence of a range of **simple** language structures **little or no attempt** to use more complex forms - grammatical errors which **impede** communication	**adequate range** of lexis and phrases - uses lexis and phrases which may be **inappropriate** to task purpose - lexical errors which **impede** communication
N/A		0	Response does not meet all the positive descriptors in B1.1 **OR** task not attempted **OR** 50% or more of the response is irrelevant			

Cap on all criteria (response length)	**Task fulfilment cap** (use of prompts)
161–219 words: **B2.2** 91–160 words: **B2.1** 71–90 words: **B1.2** 0–70 words: **B1.1**	Includes ideas from fewer than two prompts: **B2.2**

20 Appendix 7 – Oxford Test of English Advanced Writing Script 2 (Summary) marking criteria

CEFR		Marks	Task fulfilment	Organization	Grammar	Lexis
C2	C2.1	7	• response skilfully **redrafts** the main ideas with appropriate supporting details from both texts • response is **consistently clear**, **sophisticated** and appears **effortless**, with **minimal** redundancy • register is **consistently appropriate** for task purpose	• **reconstructs** ideas to produce a response with a natural flow • **consistently** coherent; well structured with **logical sequencing** of ideas • uses **sophisticated cohesive features** appropriately at all times	• **exploits** grammatical resources **creatively** to write with a distinct voice • maintains **consistent** grammatical control to produce a **very concise** response • maintains a **high level of accuracy** throughout; errors are rare and only concern complex forms	• **exploits** lexical resources **creatively** with a high degree of sophistication • maintains **consistent** control of lexis and phrases to produce a **very concise** response • maintains a **high level of accuracy** of both lexis and phrases; errors are rare and difficult to spot
C1	C1.2	6	Meets the positive descriptors in C1.1 comfortably, and the negative descriptors in C1.1 minimally if at all			
	C1.1	5	• response **synthesizes** the main ideas with appropriate supporting details from both texts • response is **clearly communicated** with **little** redundancy • register is **nearly always appropriate** for task purpose	• **reorganizes** ideas in a logically connected way • consistently coherent; **well-organized progression** of ideas • uses **appropriate cohesive features** with rare instances of misuse	• **exploits** grammatical resources to **adapt** grammatical structures • maintains grammatical control to produce a **concise** response • maintains a **good level of accuracy**; occasional errors when **adapting** grammatical structures	• **exploits** lexical resources to **adapt** lexis • **maintains** lexical control to produce a **concise** response • maintains a **good level of accuracy**; occasional errors when **adapting** lexis
B2	B2.2	4	Meets the positive descriptors in B2.1 comfortably, and the negative descriptors in B2.1 minimally if at all			
	B2.1	3	• response **synthesizes** at least two main ideas with some supporting details from both texts • response is **generally clearly communicated**; shows awareness of task purpose • register is **generally appropriate**; response shows awareness of task purpose	• **integrates** ideas appropriately; some manipulation of the order of ideas • **generally coherent**; able to connect ideas across sentences • uses **simple cohesive features** to link sentences, generally appropriately	• uses grammatical resources to **paraphrase** some grammatical structures • **moderate** grammatical control; response may **lack conciseness** • **generally** accurate; grammatical errors occasionally impede communication	• **uses** lexical resources to **paraphrase** some words and phrases • **moderate** lexical control; response may **lack conciseness** • **generally** accurate; lexical errors occasionally impede communication
B1	B1.2	2	Meets the positive descriptors in B1.1 comfortably, and the negative descriptors in B1.1 minimally if at all			
	B1.1	1	• response **includes** at least one main idea from one text • response is **not always clearly communicated** • register is **not always appropriate**; limited awareness of task purpose	• **reproduces** ideas in their original sequence; little attempt to change the order of ideas • **not always coherent**; presents ideas as a **list** of separate points • uses **some simple cohesive features** to link sentences	• uses grammatical resources to **paraphrase** in a **simple fashion** • **limited** grammatical control; response **relies** on original structures • **sometimes inaccurate**; errors may occur when **paraphrasing** simple structures	• uses lexical resources to **paraphrase** some words in a **simple fashion** • **limited** lexical control; response **relies** on original wording • **sometimes inaccurate**; errors may occur when **paraphrasing** frequent lexis
N/A		0	Response does not meet all the positive descriptors in B1.1 **OR** task not attempted **OR** response is irrelevant (i.e. off topic)			

Cap: Use of input texts	Cap: Maximum word count

- Only **one** input text is used: **B1.2** (for **Task fulfilment** <u>and</u> **Organization**)	The word limit for this task is 100. Therefore:
	    - Up to 105 words: any mark may be awarded
	    - 105–120 words: **B2.2** (for **all** criteria)
	    - 121 words or more: **B1.2** (for **all** criteria)

21 Appendix 8 – Speaking module sample responses, marks and commentaries

Below are examples of test taker responses at different CEFR levels for Speaking scripts and marks awarded. Script 1 consists of Part 1 (Interview), Part 2 (Voice message), and Part 3 (Summary). Script 2 consists of Part 4 (Debate) and Part 5 (Follow-up questions). See Appendix 5 for Speaking marking criteria.

21.1 Speaking Script 1 – example 1

This is an example of a C1-level response.

Part 1 – Interview:
Question 1. Thinking about celebrations, what was the last thing you celebrated?
'The last thing I celebrated were, erm, family birthdays, erm, many people's, many people's birthdays fall on similar dates, so I celebrated that.'

Question 2: How important is it for friends to share the same interests?
'I think it's rather important, erm, because otherwise you wouldn't have, erm, you wouldn't have common things you like to do and then to … then it would be just very boring.'

Question 3: If you could travel anywhere in the world, where would you go?
'I would really like to, erm, travel to the United States. I'm interested in many different cities over there, erm, and I would like to see, erm, for example, Los Angeles and, erm, and also San Francisco in California. Erm, and there are many beautiful natural landmarks also.'

Question 4: Finally, who has had the greatest influence on your life?
'Erm, many people had big influences on my life, but, erm, the person I'd like to mention the most is my grandmother, erm, with whom I spent lots of time and she's taught me many things and … she's always been great.'

Part 2 – Voice message:
You study at college. You had arranged to meet your tutor today to discuss something important but you now have to cancel the meeting. Leave a voice message for your tutor and:

• give a reason for cancelling the meeting

• say how you feel about cancelling the meeting

• explain why it is important to rearrange the meeting quickly.

'Good afternoon, Mister. Erm, I'm, I'm sorry to call you, but erm I'm, but I will need to cancel our, our, er, arranged meeting because unexpected test has been scheduled. I'm, I'm very so-, disappointed that that I won't be able to attend but, er, but it- but it's mandatory, but it would be very important to reschedule soon. So please call me back.'

Part 3 – Summary:
Your tutor has asked you to summarize some research for your tutor group.
Listen to two experts talking about achieving ambitions. The two experts make the same two main points. You should:
- *combine the information from the two experts and*
- *summarize the two main points the experts make.*

Expert 1: There is evidence to show that people who set ambitious goals for themselves tend to be more content than those who have lower expectations. In a recent study, participants were asked to complete a puzzle. They chose between an easy or a difficult version, but in fact they all received the same puzzle. Interestingly, those who thought they had done a more difficult puzzle reported greater pleasure at having completed it. While attempting to do something challenging brings satisfaction, this is only the case if your goal is achievable. You might, for example, decide to run a marathon in six months' time having never run before. If, after six months, you are capable of running 10 kilometres, you might view this as a failure despite other people thinking the opposite.

Expert 2: Knowing what your goals are and giving yourself challenges will make you a happier person in the long run. Without goals, life may seem meaningless: going about your daily activities, responding to other people's demands, without any sense of purpose. Too often goals are influenced by what others do; 'She can swim 2 kilometres – I could do that too!', but what's important is being able to actually reach the goals you set for yourself. And, this is where the problem with goal setting lies. It's about stretching yourself and aiming high, while accepting the reality of what is in fact possible.

Task-specific marking guide:

Main point	*E.g. mention or paraphrase of one of the following*
Main point 1: *Setting ambitious goals is linked to happiness*	*Setting ambitious goals/giving yourself challenges/knowing what your goals are makes you happier/more content*
Main point 2: *Importance of being able to achieve goals*	*Attempting something challenging/the goals you set for yourself will bring satisfaction only if the goal is achievable/you can reach the goal*

'The experts talked about, er, how important it is to, er, set challenging but, er, but also doable, er, goals for yourself. The first main point was that setting higher, er, more dif-, more, seemingly more difficult goals for yourself, er, result in more happiness and better satisfaction in the end, when you achieve your goals. But the second point was that, er, setting impossible tasks are terrible because, er, you will feel disappointed when you don't manage to hit that extremely high bar. So, er, basically they said to aim high, but stay realistic.'

Marks

Task fulfilment	Pronunciation & fluency	Grammar	Lexis
5	5	6	6

Commentary

Criterion	Mark	Comment
Task fulfilment	5	The test taker meets **the vast majority of task requirements** and responds to all parts without any evidence of misunderstanding the prompts, providing sufficient amount of information for the listener to be **fully informed within the context of the task**. In Part 1, the test taker manages

to respond comprehensively to the prompts although some of the responses could be expanded more using the allocated time. In Part 2, the first two prompts are addressed effectively while the third prompt is only re-stated, with no explanation offered. The test taker shows awareness of the diplomatic use of the language by acknowledging the situation of the listener ('I'm sorry to call you … '), expressing disappointment and stating the importance of rescheduling appropriately (' … it would be very important to reschedule soon'), however, the response could benefit from more sophistication for a completely diplomatic response. In Part 3, the test taker **synthesizes the two main points** as outlined in the task-specific marking notes and summarizes the points with a satisfactory amount of information.

The **register is generally appropriate to purpose of task and audience**. In Part 2, the test taker uses the correct level of formality to address the receiver of the voice message ('Good afternoon, Mister … ') and delivers the main points in Part 3 with an appropriate register.

Overall, the response meets the positive C1 descriptors for Task fulfilment, but only minimally due to the lack of expansion in Parts 1 and 2 and not displaying diplomatic use of language comfortably enough in Part 2.

Pronunciation and fluency	5	The response is delivered with a **fairly smooth flow**, and it is easy to follow and comprehend by the listener. The test taker generally shows **control of stress, rhythm and intonation to convey meaning**. There is no evidence of the test taker's L1 impacting on their ability to express themselves fluently. Additionally, the test taker consistently produces well-structured discourse in all parts, using **a variety of cohesive features** appropriately. However, there are numerous hesitations throughout the response, particularly in Parts 2 and 3 ('<u>Erm, I'm, I'm</u> sorry to call you, <u>but erm</u> <u>I'm, but</u> I will need to cancel <u>our, our, er,</u> arranged meeting … '), which suggests that their ability to express themselves spontaneously is not sufficient to comfortably meet the C1 descriptors for Pronunciation and fluency.
Grammar	6	The response maintains **a high level of grammatical control of both simple and complex structures**. Examples of complex structures can be found throughout the response such as adverbial clauses ('otherwise you wouldn't have, erm, you wouldn't have common things'), conditional statements ('I would really like to, erm, travel to the United States'), and fronted noun phrases ('the person I'd like to mention the most is …'). Overall, the test taker maintains high accuracy in their response and does not make many grammatical mistakes, meaning that **errors are rare and difficult to spot**. Therefore, the response meets the positive C1 descriptors for Grammar comfortably. This is close to 7 marks but the range of grammar in the response is not sufficiently wide to demonstrate mastery of a very wide range of structures.
Lexis	6	The test taker uses **a broad lexical repertoire** relevant to the tasks with a good range of frequent and less frequent lexis. There is **little evidence of the test taker needing to use search and avoidance strategies**. When responding to the tasks, the test taker shows awareness of **collocations** such as 'fall on similar dates' and 'hit that extremely high bar'. These

		expressions are used naturally and effectively to respond to the task. Therefore, the response comfortably meets the C1 descriptors for Lexis.

21.2 Speaking Script 1 – example 2

This is an example of a B2-level response.

Part 1 – Interview:

Question 1: Thinking about daily life, do you prefer spending time by yourself or with other people?

'I can be considered as really [?], so I … it depends. I normally prefer to, er, spend time alone, er, and have some 'me' time, but it's also great, er, to stay with your family and friends when you want to disconnect and it's right also.'

Question 2: Tell me about an interesting place that you have visited.

'Er, speaking for myself, the most interesting place I have visit is Vienna. I really like it. Er, it's an incredible place for have a taste of Europe past, er, while you come across, er, the … er, the most beautiful streets in the world.'

Question 3: How important is sport to you?

'Er, I really keen on sport. I practice football, er, and I really like it, er, and I also, er, can be considered as a big fan of Real Madrid team. Er, I-, when I feel down in the dumps, I see the matches and I, er, go to the top.'

Question 4: Finally, if you had to live without a mobile phone or a computer, which would you choose to live without?

'Er, I probably use more my mobile phone, so I prefer to live without a computer. Er, it's also necessary for my works in school, but, er, living without my mobile phone, er, will be a disaster for me, so I probably choose that option.'

Part 2 – Voice message:

You study at college. Another student's laptop has broken and he has asked to borrow your laptop for a few days to complete an important assignment. You do not want this student to borrow your laptop. Leave a voice message for the student and:

• *say why you are calling*

• *explain why you do not want to lend your laptop*

• *make a suggestion about what he could do.*

'Talking about if I am going to borrow you my laptop, I finally decide not. Because, er, in the last days, I have recently borrow you my mobile phone one time, and now it's broken. So I … no going to borrow you my laptop. Er, it-, it's completely necessary to use laptop. You can probably ask other people or go to the library.'

Part 3 – Summary:

Your tutor has asked you to summarize some research for your tutor group.

Listen to the two experts talking about stress. The two experts make the same two main points.

You should:

• *combine the information from the two experts and*

• *summarize the two main points the experts make.*

Expert 1: Stress is a series of physiological changes that enable us to fight back or run away when a threat is present. Most people view stress as something negative, but in reality, when stress is relatively mild and short-lived, for example, the stress experienced immediately before giving a talk to

an audience, it improves our ability to think clearly, to remember things and pay attention. In this way, controlled, manageable stress can be beneficial. On the other hand, when people feel stressed over long periods of time, often as a result of too much pressure at work or difficulties in other areas of their lives, it can give rise to negative physical symptoms such as tiredness or high blood pressure.

Expert 2: We all know that feeling we get when suddenly the pressure's on and we need to perform. When we experience stress like this, it causes our brain cells to increase in number – with communication between them improved, which helps us to work more efficiently. This natural response to stress helps us to deal with challenging tasks. However, consistently high stress levels generally lead to health problems. Because today's fast pace of living generally requires people to carefully balance multiple responsibilities associated with work and personal life – which might contribute to stress – it's important to take time to relax in order to reduce its impact.

Task-specific marking guide:

Main point	E.g. mention or paraphrase of one of the following
Main point 1: *Stress can be a good* *thing when temporary and in* *small amounts.*	*When stress is relatively mild and short-lived … it improves our ability to think clearly/to remember things /pay attention* *Controlled, manageable stress can be beneficial* *[Stress] causes our brain cells to increase in number – with communication between them improved* *[Stress] helps us to deal with challenging tasks.*
Main point 2: *Chronic stress is bad* *for people's health.*	*When people feel stressed over long periods of time … it can give rise to negative physical symptoms* *Consistently high stress levels generally lead to health problems.*

'Well, talking about the stress, er, is generally accepted that the stress is always something bad. But, er, after listening to some experts, er, stress, er, can be considered as something positive if it's over-, it is isn't over a long period of time. Er, because it, for example, can improve, er, our performance in really difficult situations. But if this stress is not controlled, can-, it can lead to huge health problems that can, er, turn our lives in-, into a really difficult life.'

Marks

Task fulfilment	Pronunciation & fluency	Grammar	Lexis
3	3	3	3

Commentary

Criterion	Mark	Comment
Task fulfilment	3	The test taker **generally fulfils the task requirements,** and the **listener is adequately informed within the context of the task**. In Part 1, the test taker understands the questions and responds with reasons and examples, but points made are not always sufficiently expanded, e.g. Question 3 requires the test taker to say *how important* sport is, but the test taker

		does not do this. The test taker performs well in Part 3 and **paraphrases the two main points** that are specified in the task-specific marking guide ('stress, er, can be considered as something positive if it's over- … if this stress is not controlled, can-, it can lead to huge health problems') and also touches upon the supporting details, reflecting the marking guide (' … it, for example, can improve, er, our performance in really difficult situations'). However, in Part 2, the test taker does not fulfil the task requirement outlined in the rubric to be *diplomatic* and appears abrupt ('I finally decide not', 'So I … don,t going to borrow you my laptop').
		The test taker covers all three prompts but does not approach the task as if they were leaving a voice message (i.e. there is no greeting or valediction). Therefore, while **register is generally appropriate to purpose of the task and audience** in Parts 1 and 3, in Part 2 it is not. Overall, the response only minimally meets the positive descriptors of B2.
Pronunciation and fluency	3	The response is generally intelligible and easily followed. The test taker **employs stress rhythm and intonation to support the meaning**, but **influence from other language(s) is evident**, which occasionally affects the ease of comprehension ('I can be considered as really [?]'; /fɪl/ instead of /fiːl/). The test taker **expresses themselves in an even tempo**, lacking spontaneity and smooth transitions. However, the test taker manages to produce **a coherent discourse, using a limited variety of cohesive features** ('so', 'also', 'because'). Overall, the response minimally meets the positive descriptors of B2 level.
Grammar	3	The test taker maintains **control of adequate range of structures** to convey the message, although **noticeable errors occur** ('I really keen on sport', 'I finally decide not'). There are frequent **attempts to use complex grammatical forms** to express viewpoints ('… it can lead to huge health problems that can, er, turn our lives in-, into a really difficult life', 'the most interesting place I have visit is Vienna') which rarely result in **mistakes that might impede understanding** ('… it's an incredible place for have a taste of Europe past'). Overall, the response meets the positive descriptors of B2, but only minimally due to the frequency of errors.
Lexis	3	The test taker uses an **adequate range of lexis to express viewpoints on the everyday topics**; however, **noticeable errors occur and may cause some confusion** ('I, er, go to the top', 'I am going to borrow you my laptop'). The test taker uses some **appropriate collocations and expressions** ('have some 'me' time', 'feel down in the dumps'), but **lexical gaps cause hesitation** ('while you come across, er, the … er, the most beautiful … '). Therefore, the response barely meets the positive descriptors of B2.

21.3 Speaking Script 2 – example 1

This is an example of a C1-level response.

Part 4 – Debate:

Most countries have a five-day working week. This should be reduced.

- *impact on the economy*
- *improved mental and physical health*
- *reduced costs*
- *customer satisfaction*
- *not suitable for all industries*

'So regarding the topic that, erm, most countries have five days of working week, but this should be reduced, I do agree that, erm, in, erm, some countries, erm, especially countries that have long, erm, that have jobs that have a lot of, erm, mental, erm, stress and also, erm, could be physical challenging, erm, a four days working week can be beneficial. Most people are not productive all the time, erm, they work, er, during the working hours, er, but productivity and concentration changes so, erm, I don't think, er, that it would erm impact, er, the economy, the productivity of the industries, because, erm, people will always do the same, but just in less time, and that could improve mental health and physical health of the workers. Er, it could also reduce costs, as, erm, you may have, erm, need a smaller working space, but not necessarily, erm, have less people working because you could have turns and, er, even possibly employing more people. So, erm, increasing the number of employed people in a in a country.'

Part 5 – Follow-up questions:

Question 1: The topic of your debate was the amount of time people spend working. If people were given more free time, do you think they would spend it wisely?

'Not necessarily, erm, but they would have the possibility to spend the time more wisely, to spend more time with their families or doing sports, erm, work on, erm, side projects that might erm be able to increase their work, erm, during working hours. For example, someone that has a non-creative job, but has quite creative abilities might have the time to work on a creative project on the side, increasing the o- overall happiness of the person.'

Question 2: What is more important; a good salary or more time off work?

'I think … probably and personally more time off work. Of course, with a balance you still need, erm, salary that allows you to live, but doesn't really make any sense to have a good salary and no time or possibility to spend it. Erm, so you could have a long erm a good salary er and short time only for a small period, but erm in a long period time working, more time, more freedom is better.'

Question 3: One day robots might do most of the work that humans do. Would this be a good thing?

'Erm, it could be helpful. I'm not sure it could be a good thing. There will still be people needing personal contact, a lot of work would still need someone physically to do the work, er, also because on the other side, the customer may want to have someone to speak to instead of a robot, but it could help with some process. It already did in the past and could definitely help to have some more general free time.'

Question 4: Some people say that it's important to have one clear ambition in life. Do you agree?

'Well, I think those people probably do have a clear ambition in life. Er, most people don't. And ambition can change during your lifetime. Also, if you have one clear ambition and you don't reach it, you won't be satisfied at the end of your life, while if you have smaller ambition or several different one, it might be easier for you to achieve some of them … erm, and be generally more happy with your life, so it's good to have dream, but it's still difficult sometimes to get to it.'

Marks

Task fulfilment	Pronunciation & fluency	Grammar	Lexis
5	5	5	5

Commentary

Criterion	Mark	Comment
Task fulfilment	5	The test taker **mostly fulfils the task requirements** and responds to both parts with generally well-developed ideas, delivered appropriately. In Part 4, the test taker's position on the debate statement is clear and a consistent argument is sustained throughout, using three of the ideas presented by the task (*improved mental and physical health*, *impact on the economy*, *reduced costs*). However, a concluding sentence is not present. In Part 5, the test taker provides sufficient amount of information relevant to the questions. Across the script, points made are sufficiently expanded, meaning the **listener is fully informed within the context of the task**, although the response to Part 4 would benefit from more explanation as to why or how mental and physical health would be improved. In both parts, **register is appropriate to purpose of task and audience** with appropriate level of formality and choice of language used ('So regarding the topic', ' … side projects that might erm be able to increase … '). Overall, the script meets the C1 descriptors for Task fulfilment but not comfortably.
Pronunciation and fluency	5	The response is delivered with **a fairly smooth flow**, and it is easy for the listener to follow and comprehend. The test taker effectively **controls stress, rhythm and** intonation to convey meaning. There are **occasional lapses**, which **do not affect intelligibility or effectiveness**. There is no evidence of the test taker's L1 impacting on their ability to express themselves fluently, although some hesitation is noticeable, especially in Part 4. Additionally, the test taker generally produces well-structured discourse across the script, using **a variety of cohesive features** appropriately ('if you have smaller ambition or several different <u>one</u>', ' … regarding the topic'). However, the organization of Part 4 could have been improved by more effective use of discourse markers to improve the flow and ease the comprehension of the message. Overall, the script meets the positive C1 descriptors for Pronunciation and fluency only minimally.
Grammar	5	The test taker uses **a wide range of structures** and generally maintains **a high level of grammatical control** of both simple and complex structures ('… someone that has a non-creative job, but has quite creative abilities might have the time to work on a creative project on the side, increasing the o- overall happiness …'). Although **errors are not absent,** they are **non-impeding** and generally occur when articulating complex thoughts ('… because you could have turns and, er, even possibly employing more people.'). Overall, the script meets the positive C1 descriptors for Grammar, but not comfortably.

Lexis	5	The test taker uses **a broad lexical repertoire** relevant to the task ('productivity', 'concentration', 'balance'). The test taker uses **a good range of collocations** appropriately ('working space', 'side projects'). Across the script, there is **little evidence of searching and avoidance strategies**. However, **errors are present** and **easy to spot** but they do not cause confusion ('physical challenging', 'small period', 'on the other side'). Overall, this script meets the positive C1 descriptors for Lexis, but only minimally due to the number of lexical errors.

21.4 Speaking Script 2 – example 2

This is an example of a B2-level response.

Part 4 – Debate:

Most countries have a five-day working week. This should be reduced.

- *impact on the economy*
- *improved mental and physical health*
- *reduced costs*
- *customer satisfaction*
- *not suitable for all industries*

'I do not think this is a good idea. While there's definitely pros and cons about this, I'm thinking more about the cons … because they are the more. First of all, the impact on the economy, er, would be huge, er, since … if there are already two free days and we make it bigger. Then, the money is not moving that much. Second of all, mm, while yes, there may be improved mental physical health, it all revolves around, basically, how much money you have. Even if they are more than two, three days a week, if a person does not have enough money, they're gonna keep working if it's another work. And it's basically gonna be the same. So, no. Er, the customer satisfaction is another point … in areas like in pubs and like that. Mm, closing more than two time- two days a week would be disastrous ... because the time the people go out more is on weekends and that's the time when people want to be free. That's why it's not really suitable for in a situation like that like that and it would have a huge impact on how much the pubs and such make, lowering … erm ...'

Part 5 – Follow-up questions:

Question 1: The topic of your debate was the amount of time people spend working. If people were given more free time, do you think they would spend it wisely?

'I do not think so, at least for the healthy class people, because they already have so much free time. What are they gonna do with more but with working class people, as I said before, maybe if they have so much free time, they're gonna, mm, search for a side job or something like that. And they're not actually gonna spend that time on enjoying it on another things or like that because it's something out of necessity.'

Question 2: What is more important; a good salary or more time off work?

'I think it's more important like good salary. Not only because of the health issues it would come with more work than we already have. But because with that money, you can / can't [?] afford time. Time is the key. If you have time, you can do the things you enjoy, and your health doesn't … go mentally bad. And you just enjoy it.'

Question 3: One day robots might do most of the work that humans do. Would this be a good thing?

'This isn't inherently a good or bad thing, because I think by the time the robots are gonna be doing most of our actual jobs, there's gonna be even more jobs for humans to occupate. So I don't think this is a actually bad thing. It might be good because it would lower the stress a lot with, erm, working class people since they not- don't have that pressure.'

Question 4: Some people say that it's important to have one clear ambition in life. Do you agree?

'On this topic, I agree a lot, because if you have a motive to be in life, then you are more … excited to do things in order to … get this … And also you work harder. So you can accomplish it with you don't. If you don't really know what to do, erm all in all, I think it's something all people should have.'

Marks

Task fulfilment	Pronunciation & fluency	Grammar	Lexis
3	4	4	4

Commentary

Criterion	Mark	Comment
Task fulfilment	3	The test taker **generally fulfils the task requirements**, and the **listener is adequately informed** within the context of the task. For the Part 4 Debate, the test taker's position is fairly clear and consistent throughout, referring to three of the ideas presented by the task (*impact on the economy, improved mental and physical health* and *customer satisfaction*). Although the test taker provides relevant information and maintains their train of thought, points made are rarely expanded or not appropriate to support their case (*improve mental and physical health*). Besides, a concluding sentence is not present. In Part 5, the test taker understands all questions and provides relevant answers generally with an adequate amount of information, however there are occasions when viewpoints are not clear and consistent (Part 5 Question 2). **Register is generally appropriate** to purpose of task and audience ('it might be good …', '… since they don't have …') although some language forms and phrases do not reflect the appropriate level of formality ('gonna', 'something like that'). Overall, the script meets the B2 descriptors for Task fulfilment, but the positive descriptors are not met comfortably so it is awarded 3 marks.
Pronunciation and fluency	4	The response is generally intelligible and easily followed. The test taker generally **employs stress rhythm and intonation** to support the meaning. Although there is an instance where pronunciation may cause confusion (' … because with that money, you can't / can [?] afford time'), that does not prevent the script from meeting the positive B2 descriptors comfortably across the parts. The test taker shows an ability to express themselves in **stretches of language with an even tempo**. A good **variety of cohesive features** are used to provide generally organized responses ('First of all', ' … is another point', 'That's why … '). Some hesitation is noticeable, but **lapses do not affect meaning**. Overall, the script meets the positive B2 descriptors for Pronunciation and fluency comfortably and the negative descriptors only minimally.
Grammar	4	The test taker uses **an adequate range of structures** to convey the message, generally **maintaining control in simple and complex structures** ('while yes, there may be improved mental … ', 'closing more than two time- two days a week would be disastrous'). However, some errors are also present ('because they are the more', 'if there are already two free days and we make it

		bigger, then, the money is not moving that much'). Overall, the script comfortably meets the positive C1 descriptors for Grammar, but the level of control and the variety of structures are not sufficient to be awarded a C1.
Lexis	4	The test taker uses **an adequate range of lexis** to express viewpoints with some **appropriate collocations**. Although the majority of the lexis in the response is fairly high-frequency there are occasional uses of less frequent lexis ('pros and cons', 'side job', 'accomplish', 'revolves around') which elevates the response. However, errors are present and frequent use of vague phrases to fill in **lexical gaps** is noticeable ('something like that', 'another things or like that'). Overall, the response is sufficient to meet the positive B2 descriptors for Lexis comfortably, but the lexical gaps and errors hold it back from being awarded a higher mark.

22 Appendix 9 – Writing module sample responses, marks and commentaries

See Appendices 6 and 7 for Writing marking criteria (Script 1 (Essay) and Script 2 (Summary) respectively).

The Script 1 samples are responses to the following task:

The clock shows how much time you have to write an essay. Write 220–280 words.

You have been discussing the topic of social media in your class. Your tutor has asked you to write an essay on the following:

> The advantages of social media outweigh the disadvantages.
>
> Do you agree or disagree?

Your essay must include at least **two** of the following ideas:

- impact on health
- impact on learning
- impact on relationships.

Organize your essay clearly, introducing the topic, providing support for the points you make, and giving a conclusion.

Write your **essay**.

Write your answer here …

22.1 Writing Script 1: Essay – example 1

This is an example of a C1-level response.

Social media has become a huge part of our lives in just a few decades and has had major consequences on our lives and on our behaviour. I believe that unfortunately its disadvantages are greater than its advantages, and in this essay I will detail my reasons doing so.

Firstly, there are some advantages to the fact that social media platforms have become so popular, one of which is that it has created many jobs. For example marketing and PR departments now put a wast amount of time and effort into their social media presence, and advertising on these sites is more popular than ever.

On the other hand, social media has many negative impacts that also need to be considered. One of them is that images can be edited, a distorted reality can be shown by ordinary people, and thus can create a false reality with a false sense of body images. This sadly has lead to a noteable rise in body image issues and eating disorders, especially among teenage girls, who are both the extremely impressionable and use these platforms very often. Another reason against social media is how it distrots the way we communicate. Texting does not only exist on social media sites, but they have helped in the rise of their popularity. This alone would not be a negative thing, but it has changed the way relationships are in the 21st century, mostly for the worst.

All in all, social media is very complex and has both advantages and disadvantages, however sadly the latter outweight the former, because of the effects on our minds, bodies, and relationships.

[271 words]

Marks

Task fulfilment	Organization	Grammar	Lexis
5	5	5	5

Commentary

Criterion	Mark	Comment
All (caps)	N/A	No caps are applied.
Task fulfilment	5	This is a **well-developed** essay with a clear and consistent argument running through it from introduction to conclusion. Points are **expanded** with relevant examples and justification, meaning the reader is generally not left having to fill in gaps, and the essay is well structured, making the argument easy to follow. **Overall the impact on the reader is consistently positive**. The impact on the reader would have been even more positive if the grammatical and lexical errors had been fewer, and if all of the points had been expanded to their full potential: for example, the second paragraph would have benefited from a fuller explanation as to how social media creates jobs and why this is a good thing; and the third paragraph is missing an explanation as to why texting has changed relationships for the worse. As it stands, the essay minimally meets the C1 descriptors.
Organization	5	The ideas are **well organized** and **the text is consistently coherent and well structured**. There is a clear introduction in which the test taker states their opinion in response to the essay question. The main body paragraphs mirror the two main elements in the essay statement, with the first paragraph covering the advantages and the second the disadvantages of social media. The conclusion is consistent with the previous paragraphs and is justified with an effective summary of the disadvantages that the test taker has mentioned in the main body ('the effects on our minds, bodies and relationships'). While the test taker uses **a range of cohesive features appropriately**, a lack of sophistication is sometimes apparent in their use of discourse markers: 'Firstly', 'On the other hand', and 'All in all' are not used incorrectly, but they are blunt ways to link the paragraphs and fail to convey any of the precision we might expect at C1. Therefore the essay only minimally meets the C1 descriptors.
Grammar	5	The test taker uses **a wide range of structures** with **a high level of grammatical control**. For example, the sentence 'Firstly, there are some advantages to the fact that social media platforms have become so popular, one of which is that it has created many jobs' demonstrates the test taker's ability to construct multi-clausal complex sentences to express their ideas fully while maintaining their control of syntax and verb forms. The test taker **maintains a good level of grammatical accuracy throughout** and **errors are rare and difficult to spot**, though they are not rare enough for the C1 descriptors to be met comfortably.

| Lexis | 5 | The test taker **uses a wide range of lexis and phrases** which are **appropriate** to an essay on social media ('platforms', 'social media presence', 'advertising', 'edited', 'distorted reality', 'body image issues', 'eating disorders' 'impressionable'). There is a **good level of accuracy** and **errors are rare** with only one being impeding ('wast' for *vast*). However, there are enough errors to mean that the essay only minimally meets the C1 descriptors. |

22.2 Writing Script 1: Essay – example 2

This is an example of a B2-level response.

Nowadays, social media play a very big role in one's life. Some people may argue about the impact it has on us. So how it really affects us?

It is a well-known fact that people, have become more lazy ever since social media appeared. When you are on your phone all the time just scrolling, you are not exactly making an effort to improve your health. Deranged sleep schedule, doing fewer exercises, obesity are the problems that are caused because people would rather stare at their screens than do something beneficial for their health. So in this case I can see why people are concerned about phone addictions.

On the other hand, people are now learning about variety of different things much faster all thanks to social media. Eventhough, not everything on them will be appealing to us, we can still find so many interesting topics and learn about things we are keen on using social media. And the only thing we need is internet! In this field people are benefiting from them.

Same can not be said about relationships. Due to being overwhelmed by their phones and information on social media, people seem to have lost their ability to communicate in person. Chatting online appears to be easy, but when you meet with your friend in real life, it can often get akward since you do not know how to start a conversation. Social media has definitely had a negative impact on our communication scills, therefore our relationships too.

To sum up, I think there are more distadvantages than advantages of social media, but i can not argue-we would not be able to function normally without them.

[278 words]

Marks

Task fulfilment	Organization	Grammar	Lexis
4	4	4	4

Commentary

Criterion	Mark	Comment
All (caps)	N/A	No caps are applied.
Task fulfilment	4	The essay is **generally well developed**, though organizational weaknesses mean that the direction of the test taker's argument

(whether they agree or disagree with the essay statement) is not clear until the conclusion. Points are **reasonably expanded** but they don't have the precision that would be expected at C1: those addressing health and learning relate to the Internet and phones in general, rather than social media specifically. The **register is generally appropriate**, though the test taker occasionally adopts an inappropriately informal tone ('you are not exactly making an effort', 'all thanks to social media', 'Same can not be said about relationships'). Despite its faults, the essay is easy to follow and comfortably meets the B2 descriptors.

Organization	4	The organization of ideas is **generally good** with separate paragraphs for the introduction, the conclusion and the three points made in the main body. However, the different paragraphs are linked rather loosely and do not present points in a logical order (disadvantage, advantage, disadvantage). The lack of an outline of the test taker's argument in the introduction also contributes to a directionless feel to the essay. Despite this, the essay is cohesive and coherent with only one major **break in communication** ('but i can not argue-we would not be able to function normally without them'), and it comfortably meets the B2 descriptors.
Grammar	4	The test taker shows they can combine **a wide range of structures** to form well-controlled complex sentences that enable them to **express viewpoints clearly and fully**. For example, in the following sentence the test taker uses a causal clause and a comparison clause within a relative clause: 'Deranged sleep schedule, doing fewer exercises, obesity are the problems that are caused because people would rather stare at their screens than do something beneficial for their health.' However, though the test taker shows they can maintain a good level of accuracy in even the most complex structures, they still make a number of relatively basic errors in their essay, including misplaced commas, missing conjunctions and an incorrect question form ('So how it really affects us?'). These errors are not **difficult to spot** and mean that the essay does not meet the C1 descriptors. However, it does meet the B2 descriptors comfortably.
Lexis	4	The test taker uses **a good range of lexis** which is **appropriate to task purpose** ('scrolling', 'obesity', 'stare at their screens', 'beneficial', 'addictions', 'function'), though lexical limitations are apparent in the fact that their points mainly relate to Internet and phone usage in general, rather than being targeted specifically at social media. Lexis is sometimes **used unnaturally or inaccurately** ('deranged sleep schedule', 'In this field') and errors can **occasionally impede communication** ('scills'), but generally the accuracy is high. The essay does not demonstrate the lexical range needed for C1, but it does meet the B2 descriptors comfortably.

22.3 Writing Script 2: Summary – example 1

Textbook extract	Lecture transcript
One of the biggest threats to survival for most animals is the threat of being eaten by another animal. Stick insects, for example, have a number of adaptations that allow them to avoid their predators. Firstly, as their name suggests, they look just like sticks in colour and shape, which means they have very effective camouflage in their natural habitat of trees and plants. Not only does the physical appearance of stick insects make them hard to see, but so does the way they act. They tend to stay perfectly still, especially during the day, and when they do move they sometimes sway as if being blown by the wind. This means that few predators will see a stick insect, even when it is right in front of them. If a bird or other predator does manage to spot it, the stick insect may have one further trick. Some stick insects can produce nasty substances that smell terrible and prevent their predators from trying to eat them.	'So far we have looked at behavioural and structural adaptations, but there is one further way in which living things are adapted to their environment, and these are physiological adaptations. A physiological adaptation is basically a function or process inside the body that allows an organism to compete and survive. For example, the nettle is a plant that's known for its painful sting. The sting is caused by a poison that the plant produces, and this is a physiological adaptation which prevents animals eating its green leaves. Remember – don't confuse physiological with structural, which is an adaptation related to the actual appearance and shape of the organism. Behavioural adaptations are easy as it's just about how something behaves. Every plant and animal will have examples of each of these different types of adaptations.'

Task-specific marking guide

	Main idea	Supporting details
1	Every living thing has adaptations that help it to survive **(2)**	N/A
2	Behavioural **(2)**	Stick insects act like sticks for camouflage (staying still and swaying as if blown by the wind) **(1)**
3	Structural: appearance/shape **(2)**	Stick insects look like sticks for camouflage **(1)**
4	Physiological: 'function or process inside the body' **(2)**	Some stick insects produce bad-smelling substances to put off predators **(1)** Nettles have a sting which prevents animals from eating them **(2)**

Key
   (1) From the textbook extract
   (2) From the lecture transcript		

This is an example of a C1-level response.

Animals have different methods of trying to avoid predators, and stick insects are a great example to examine the three types of adaptations. Structural adaptation refers to the outside of the animal, which in this case is that stick animals look like small sticks. Behavioural adaptation concerns their actions, for example that these animals stay perfectly still or move in a way that blends into their habitat. Lastly their physiological adaption is that they omit something which causes a terrible smell to scare off potential predators. [86 words]

Marks

Task fulfilment	Organization	Grammar	Lexis
5	6	6	6

Commentary

Criterion	Mark	Comment
All (caps)	N/A	No caps are applied.
Task fulfilment	5	This is an effective summary of the information in the input texts (as outlined in the task-specific marking guide). The three types of adaptation are set out clearly and supported **appropriately** with examples. For the structural and behavioural adaptations, the test taker provides both an explanation and an example. However, when covering the third type of adaptation, physiological, the test taker does not attempt any kind of definition, and their single example is not sufficient to allow a reader to fully understand the meaning of the term. We should note, though, that physiological adaptation is particularly hard to explain as a concept, so failing to elucidate it in an otherwise **clearly communicated** response should not prevent the test taker from minimally meeting the C1 descriptors.
Organization	6	The summary presents a **well-organized progression of ideas**. The opening sentence sets out the overall main idea (adaptations are ways for animals to avoid predation) and outlines the key information around which the rest of the summary will be structured (there are three types of adaptation and they can be exemplified by stick insects). The next three sentences then cover the adaptations in turn, each sentence ending with an example showing how that adaptation relates to stick insects. By integrating the stick insect examples (which come from the textbook extract) with the descriptions of the different adaptations (which come from the lecture transcript), the test taker has **reorganized ideas in a logically connected way**. **Appropriate cohesive features** are also used throughout (reference: 'the three types of adaptations', 'their actions', 'these animals', 'their physiological adaptation is that they …'; discourse markers: 'in this case', 'for example', 'Lastly'). While the summary lacks the sophistication and **natural flow** required by the C2 descriptors, it does meet the C1 descriptors comfortably.
Grammar	6	The test taker is not at all reliant on the grammatical structures in the texts, but communicates the ideas contained within them in a clear and **concise** way using their own structures (often synthesizing from more than one place). For example, they **adapt** and shorten *when they do move they sometimes sway as if being blown by the wind* from the textbook extract by changing it to: '[they] move in a way that blends into their habitat.' The test taker also **maintains grammatical control** and **a good level of accuracy** throughout, with no clear errors. The summary does not meet the C2 descriptors as it is not **a *very* concise response** nor does it convey **a distinct voice**, but it does meet the C1 descriptors very comfortably.

Lexis	6	The test taker does not simply paraphrase words from the texts but shows their comprehension of the ideas in the way that they **concisely** re-explain some ideas in a new way. For example, they **adapt** and shorten *structural, which is an adaptation related to the actual appearance and shape of the organism* by changing it to 'Structural adaptation refers to the outside of the animal'; with the use of the phrase 'the outside', they demonstrate that they have identified the property that appearance and shape both have in common. Similarly, when they change *sway as if being blown by the wind* to 'move in a way that blends into their habitat', they show they have recognized and can explain how the movement acts as camouflage. While doing this, the test taker **maintains lexical control** and **a good level of accuracy**, making **only occasional errors**. The response does not meet the C2 descriptors as it is not *very* concise nor does it have a **high degree of sophistication**, but it does meet the C1 descriptors comfortably.

22.4 Writing Script 2: Summary – example 2

Textbook extract	Lecture transcript
Repairing products to extend their lives was once common practice, but today planned obsolescence and the complexity of electronic devices mean that it's often quicker and cheaper to buy a replacement than to attempt a repair. Even simple repairs like replacing a battery can be too costly or difficult. However, inspired by a growing design for repair movement, some companies now specialize in producing smartphones which are easy to repair or upgrade, with affordable replacement parts and clear instructions available. A design principle related to design for repair is design for disassembly which is about ensuring products can easily be taken apart at recycling centres at the end of their lives. For example, smartphones contain a range of metals (including rare ones) as well as plastic and glass. If these materials can't be separated from each other, they will be wasted, which is why design for disassembly is so important.	'Thinking about the end of a product's life is an essential design consideration. Have you ever cracked your smartphone screen or had a weak battery? Did you get it repaired? Or decide to get a new phone? Devices only last as long as the first components to fail, whether that's the battery or screen or, um, it could be the software which reaches the end of its life first. Smartphone software is typically only supported with updates for three to five years. After this, users may find they can no longer update apps and that their phones become vulnerable to security problems. When manufacturers expect products to need replacement after a relatively short life, this is known as planned obsolescence. It's attractive to them because it means consumers are forced to buy new models on a regular basis. However, the extra waste created is clearly bad for the environment.'

Task-specific marking guide

	Main idea	Supporting details
1	Planning 'the end of a product's life is an' important part of design **(2)**	N/A
2	Planned obsolescence: 'when manufacturers expect products to need	The complexity of electronic devices means 'it's often quicker and cheaper to buy a replacement than to attempt a repair' **(1)**

	replacement after a relatively short life' **(2)**	'Devices only last as long as the first components to fail', e.g. the battery, the screen or software **(2)**
		Attractive to manufacturers because 'consumers are forced to buy new models' regularly **(2)**
		The extra waste created is bad for the environment **(2)**
3	Design for repair **(1)**	Some companies produce 'smartphones which are easy to repair or upgrade' ('affordable replacement parts and clear instructions') **(1)**
4	Design for disassembly: 'products can easily be taken apart' for recycling **(1)**	Smartphones contain a range of materials that will be wasted if they can't be separated for recycling **(1)**

Key
(1) From the textbook extract
(2) From the lecture transcript

This is an example of a B2-level response.

> Nowadays, repairing items is more costly than simply discarding it and buying a replacement. However, it has a clear negative effect on the environment, since plastic, glass and metals, including those, which are considered valuable and typically not found often, are thrown away. Recently, some companies attempt to rectify this issue by making products which are easily repairable. However, certain brands do the opposite, manufacturing products which break easily, since it enables people to buy new, expensive models often. Companies do this because it is good for business. Overall, certain companies are fixing, others - utilising this issue which harms surroundings.
>
> [101 words]

Marks

Task fulfilment	Organization	Grammar	Lexis
4	4	4	4

Commentary

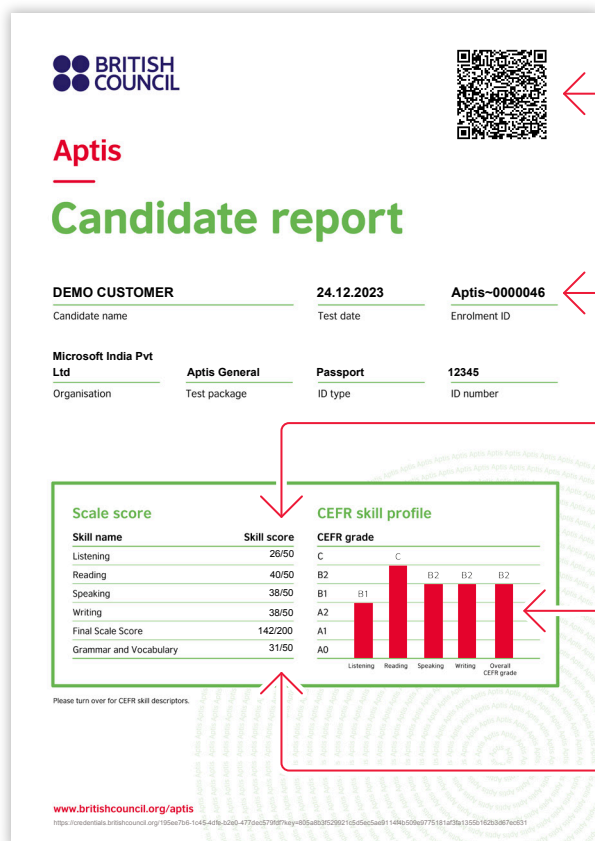| Criterion	Mark	Comment
All (caps)	N/A	No caps are applied.
Task fulfilment	4	Although the response accurately conveys a reasonable amount of information, it isn't quite successful in **synthesizing the main ideas with appropriate supporting details from both texts** as there are some important details that are missing. The test taker includes the idea that some products are designed to be repairable (Main idea 3 in the task-specific marking guide) but does not support this with information about how this might be achieved (i.e. *affordable replacement parts and clear instructions*). There is also no reference to products which are designed to be easy to recycle (Main idea 4 in the task-specific marking guide). Because

		of this, the response does not meet the C1 descriptors. However, the summary is **clearly communicated**, the register is **appropriate**, and it meets the B2 descriptors comfortably.
Organization	4	The test taker **uses appropriate cohesive features**, including reference ('this issue'), substitution ('do the opposite'), ellipsis ('certain companies are fixing, others - utilising this issue'), discourse markers ('However'), and lexical cohesion ('some companies … certain brands'). However, the response does not meet the C1 descriptors as it does not **reorganize ideas in a logically connected way**. In particular, the sentence 'Recently, some companies attempt to rectify this issue by making products which are easily repairable' has not been positioned logically, since the text before and after it relates to products which cannot easily be repaired. It would have been more logical for this sentence to appear towards the end of the summary, so that the description of the problem of lack of repairability would be more closely linked to the reference to companies who intentionally design products this way. Despite this, the test taker does show that they can **integrate ideas appropriately** with **some manipulation of the order of ideas**. This is clear from the first two sentences in which the test taker links an idea from the opening sentence of the textbook extract (*today … it's often quicker and cheaper to buy a replacement than to attempt a repair*) to a synthesis of ideas from the end of the textbook extract (*smartphones contain a range of metals (including rare ones) as well as plastic and glass*) and the end of the lecture transcript (*However, the extra waste created is clearly bad for the environment*). All in all, the response meets the B2 descriptors comfortably.
Grammar	4	The test taker **paraphrases** grammatical structures. For example, they paraphrase *it's often quicker and cheaper to buy a replacement than to attempt a repair* as 'repairing items is more costly than simply discarding it and buying a replacement' and *easy to repair* as 'easily repairable'. Note that the test taker's structures still follow those in the input texts relatively closely and don't involve the kind of transformation that we would expect at C1 (i.e. they **paraphrase** rather than **adapt grammatical structures**). However, the response does maintain **moderate grammatical control** and is **generally accurate**, with most errors being non-impeding. All in all, the summary comfortably meets the positive B2 descriptors.
Lexis	4	The test taker **paraphrases some words and phrases**, for example 'not found often' for *rare*, and 'thrown away' for *wasted*. The response is also **generally accurate** and maintains **moderate lexical control** with no clear errors. However, the test taker's use of lexis lacks conciseness. For example, they write 'it has a clear negative effect on the environment' when they could have written *it damages the environment*), and they say 'which are considered valuable and typically not found often' as a lengthy paraphrase of *rare*). This means that the response does not meet the C1 descriptors, but it does comfortably meet those at B2.

23 References

Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics, 14*(2), 115–129. https://doi.org/10.1093/applin/14.2.115

ALTE. (2011). *Manual for language test development and examining.* Council of Europe. https://rm.coe.int/manual-for-language-test-development-and-examining-for-use-with-the-ce/1680667a2b

ALTE. (2020). *Principles of good practice 2020.* https://www.alte.org/resources/Documents/ALTE%20Principles%20of%20Good%20Practice%20Online%20(Final).pdf

Bond, T., Yan, Z., & Heene, M. (2020). *Applying the Rasch model: Fundamental measurement in the human sciences* (4th ed.). Routledge.

British Council, UKALTA, EALTA, & ALTE. (2022). *Aligning language education with the CEFR: A handbook.* http://www.ealta.eu.org/documents/resources/CEFR%20alignment%20handbook.pdf

Council of Europe. (2001). *Common European Framework of Reference for languages: Learning, teaching, assessment.*

Council of Europe. (2020). *CEFR for languages: learning, teaching, assessment. Companion volume with new descriptors.* https://rm.coe.int/common-european-framework-of-reference-for-languageslearning-teaching/16809ea0d4

Finnerty, C. (2015, 01 01). The CAT is out fo the bag. *Modern English Teacher,* 24(3), pp. 15-17.

Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika, 54*(3), 427-450. https://doi.org/10.1007/BF02294627

AG_DOC_0270

BRITISH COUNCIL

Aptis

How to read the score report

Document preview (Candidate report)

BRITISH COUNCIL

Aptis

Candidate report

DEMO CUSTOMER	24.12.2023	Aptis~0000046
Candidate name	Test date	Enrolment ID

Microsoft India Pvt Ltd	Aptis General	Passport	12345
Organisation	Test package	ID type	ID number

Scale score

Skill name	Skill score
Listening	26/50
Reading	40/50
Speaking	38/50
Writing	38/50
Final Scale Score	142/200
Grammar and Vocabulary	31/50

CEFR skill profile

CEFR grade

(Bar chart: Listening B1, Reading C, Speaking B2, Writing B2, Overall CEFR grade B2)

Please turn over for CEFR skill descriptors.

www.britishcouncil.org/aptis

https://credentials.britishcouncil.org/195ea7b6-1c45-4df8-b2e0-477dec579fdf?key=805a8b3f529921c5d5ec5ae9114f4b509e977518f3a1355b162b3d67ec631

QR Code:
The validity of Aptis results can be confirmed by scanning the QR code or by following the link at the bottom of the certificate.

Test taker details and session information

Numerical Score:
Note that these scores should NOT be compared across skills. This is because benchmarks between CEFR levels vary according to skill.

CEFR Skill Profile:
Provides CEFR levels across the different skills tested. These can be compared.

The grammar and vocabulary score is only reported as a numerical score. When a candidate's score in a particular skill component falls just below a CEFR level threshold, their performance in grammar and vocabulary is taken into account in assigning the CEFR level.

How scores can be used

Numerical Score:

o To give a detailed comparison of candidate performances for a given skill within a group, including comparisons between students within the same CEFR level at a more fine-grained level.

o To track test taker performance for a given skill over a period or following language teaching/learning intervention. This is particularly relevant when intervention or learning period may not be sufficient to realise improvement over 1 or more CEFR levels.

CEFR Skill Profile:

o To provide benchmarked CEFR levels of proficiency which can be referenced to descriptions of what a language user can typically do at these levels.

o To differentiate strengths and weaknesses across skills to help provide road maps for learners and teachers to target areas for improvement (referencing the descriptions of what typical language users can do).

o To show improvement over longer periods of time or more intense interventions based on recognized criteria.
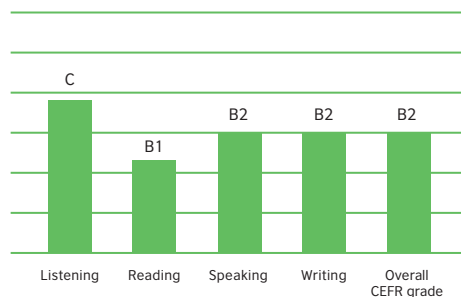
Aptis

Relationship between scale scores and CEFR skill profile

Candidate A

Overall CEFR B2

Skill Name	Skill Score
Listening	44/50
Reading	32/50
Speaking	47/50
Writing	46/50
Final Scale Score	169
Grammar and Vocabulary	30/50

CEFR Skill Profile

	Listening	Reading	Speaking	Writing	Overall CEFR grade
	C	B1	B2	B2	B2

Candidate B

Overall CEFR C

Skill Name	Skill Score
Listening	40/50
Reading	42/50
Speaking	47/50
Writing	40/50
Final Scale Score	169
Grammar and Vocabulary	35/50

CEFR Skill Profile

	Listening	Reading	Speaking	Writing	Overall CEFR grade
	C	B2	C	B2	C

Candidate A:
This candidate was awarded an overall CEFR level of B2. An overall CEFR level is only calculated for candidates who take a four-skills package.

This candidate achieved an overall scale score of 169. This is the total numerical score from the four skills.

For each language skill, and overall (if all 4 skills are taken), the candidate is awarded both a numerical scale score and a CEFR level.

Candidate A scored 30/50 in the grammar and vocabulary component.

Comparing scores:
Numerical scores can be compared for the same skills component. For example, in Listening, these two candidates have the same CEFR level, but the numerical scores inform us that Candidate A had a slightly stronger performance than Candidate B.
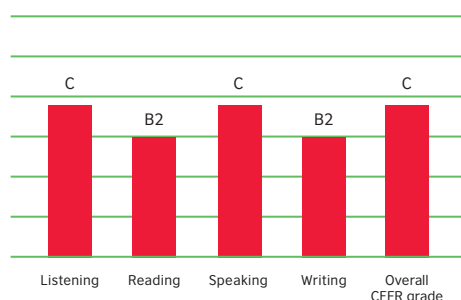
Note that numerical scores cannot be compared between different skill areas, so Listening numerical scores cannot be compared to, e.g., Speaking. This type of comparison should be at CEFR level only.

Candidate B:
This candidate received an overall CEFR level of C and an overall numerical score of 169.

Note that although Candidate A has the same overall numerical score, Candidate B receives a higher overall CEFR level. This is possible because these two summary results are calculated independently of one another. The overall numerical score is a straightforward total of the scores for each of the four skill areas (out of a maximum possible 200), the overall CEFR calculation is based on the four CEFR levels achieved.

Speaking score comparison:
Both candidates received the same numerical score for the Speaking component (47), but Candidate A was awarded a CEFR level B2 and Candidate B level C. This is not a mistake. Both candidates scored in the "grey area" just below a CEFR level threshold for Speaking. The comparatively strong performance in the grammar and vocabulary (core) component for Candidate B resulted in their being awarded level C.