GATEWAY English Test:
Test Specifications

Founding Technical Advisory Board

Ute Knoch
Director of the Language Testing Research Centre, University of Melbourne
Ph.D. Language Teaching and Learning, The University of Auckland

John De Jong
Professor Emeritus of Language Testing
Ph.D. Educational Measurement, University of Twente

Mark Reckase
University Distinguished Professor Emeritus at Michigan State University
Ph.D. Psychology, Syracuse University

Eddy White
Associate Professor, Second Language Acquisition & Teaching, University of Arizona
Ph.D. Applied Linguistics, Macquarie University

Michael Griffiths
Professor, Computer Science, Southern Virginia University
Ph.D. Instructional Technology, Brigham Young University

1 Introduction

The test specifications provide a description of the GATEWAY English test. The document sets out the rationale behind the need for a new English proficiency test, provides an overview of the test development process, its composition and the item types the test comprises. The test specifications also illustrate the scoring, delivery, and alignment of GATEWAY to the Common European Framework of Reference for Languages (CEFR).

2 Test description and rationale

Educational institutions using English as the language of instruction need valid and reliable means of assessing students' English language proficiency at key stages of their language development – preferably in relation to the widely understood levels of the Council of Europe's CEFR.

The reason for the development of this new test is the perceived need of the prospective stakeholders, especially the learners themselves, to have easy access to a high-stakes, low-cost online proficiency test for academic admissions purposes. Many institutions require information on their students' language proficiency to evaluate their readiness to pursue further studies or activities that require a specific level of English proficiency. The test thus serves the individual learner's need for external verification of their language proficiency required for study or career progression at much lower cost than comparable products.

GATEWAY has been designed to measure language proficiency most reliably at CEFR levels B1 (independent user) to C1 (proficient user). The content of the test is independent of any specific course of study and reflects the register of spoken and written academic English. It is therefore ideally suited for measuring students' proficiency in academic English at key points in their learning programs.

GATEWAY measures learners' ability to communicate effectively in English incorporating reception, production and interaction in spoken and written modes of the language. Although these are traditionally referred to as the four skills: listening, speaking, reading, and writing, when communicating in English, these four skills do not occur in isolation. Therefore, GATEWAY features a variety of integrated item types that emulate the language learners encounter in educational settings. Students are expected to combine several language skills to participate successfully in their studies.

Hence it follows that the users of test results can be confident that if the test takers' English language proficiency meets the course requirements, the test takers will be able to actively participate and succeed in their courses.

The test is delivered entirely online and can be taken on-demand.

3 Key characteristics

3.1 Designed for higher education admissions

GATEWAY was designed specifically for Higher Education admissions. The test measures students' ability to communicate effectively in English with a clear focus on academic English. Great care was taken in selecting a variety of item types that cover the domain of academic English ensuring each testing point is linked to a target language use situation. GATEWAY includes reception, production and interaction in spoken and written modes of the language that simulate the language students encounter in educational settings across academic disciplines. It features a variety of integrated item types that combine several language skills needed to actively and successfully participate on campus.

3.2 Secure

English3 provides a secure virtual testing environment in which candidates can take the GATEWAY test on demand. At English3 we are layering security measures. In general, these layers include computer-based security flagging systems coupled with human evaluation at differing levels. In particular, we utilize asynchronous video monitoring, confirmation of ID, facial recognition, plagiarism recognition software, etc. Refer to Section 8 Test Delivery for more details.

3.3 Transparent

GATEWAY provides administrators with a transparent view of test taker proficiency by including an unscored video interview and writing sample with every test. Test takers record a 10 to 15-minute interview that allows decision makers to get to know the applicant by observing their conversational abilities.

3.4 Online

All of our tests were designed for online usage, which allowed us to select item types that are valid, reliable, authentic and secure in an online environment. GATEWAY is an on-screen test that students take on their own devices at a time that suits them. The entire experience of taking GATEWAY is a seamless online process - from registration to taking the test to receiving the score report and informing the selected institutions of the test result. Each test taker has their own dashboard to view the test scores, watch their interview video and download their score report.

3.5 Affordable

Affordability is a central tenet of increasing access. An intermediate English language proficiency is the prerequisite for a successful participation in university studies. During the admissions process, applicants have to have their level of English proficiency certified. Most admissions tests are, however, prohibitively expensive for international students from less well-off economies. With

the GATEWAY test, English3 wants to ensure that more students are able to afford taking a high-stakes language exam that allows them to apply to university where English is the medium of instruction and subsequently fulfill their academic ambitions.

3.6 Accessible

Test takers complete the test at a time and place that is convenient for them. Test takers need a reliable internet connection, a laptop or computer, a microphone and speakers and a place where they can work entirely alone and undisturbed. The test must be completed in one sitting. There are two optional ten-minute breaks scheduled during the test.

4 Quality assurance

GATEWAY is produced by English3. A technical advisory group (TAG) has overseen the test development process and approved its various stages. The audits by TAG members represent a continuous process aimed at maintaining and improving the quality of the test. This involves scrutiny of the different stages of test design, production, and administration. The process will continue beyond the launch of the test to ensure that every test taker receives a fair and valid test result.

5 Test development process

GATEWAY was developed through an iterative process involving the following stages:

- ➔ Initial test design
- ➔ Drafting of test specifications
- ➔ Production of sample materials
- ➔ Technical development for online delivery, including proposed security mechanisms
- ➔ Internal and external reviews
- ➔ Modification on the basis of the reviews
- ➔ Trialing with students in teaching centers around the world
- ➔ Test production
- ➔ Field-testing
- ➔ Psychometric analysis and review
- ➔ Item banking

5.1 Test design

The first phase of test development involved producing comprehensive test specifications. The specifications detail the test format, its content and tasks. They underpin the consistency of measurement across skills sections; maintaining the quality of the test across administrations and helping to ensure that decisions based on test scores will be fair and valid.

The specifications for GATEWAY were developed in line with the Common European Framework of Reference for Languages (CEFR). Each task in the test is related to one or more CEFR scales with a focus on functional and communicative competencies.

Draft tests specifications were reviewed by an internal panel as well as the external TAG 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.

Experienced item writers were commissioned to draft item writer guidelines for each section of the test with the help of the draft specifications and sample materials. These guidelines helped the item writers to produce tasks and to ensure consistency across different instances of the test.

A team of item writers was then trained to write an initial set of test materials, which fed into small-scale trialing in which groups of students were asked to complete the trial test and provide feedback on their experience. This was followed by another round of minor revisions based on the comments from the item writers and trial students.

Further test forms were then commissioned. These were field tested more extensively on representative samples of the target population in a range of countries worldwide.

5.2 Test format

GATEWAY is delivered entirely online and contains mainly open-response and integrated item types as well as some discrete-point item types. Efforts have been made to tap into inferred or pragmatic meanings, as well as testing more concrete understanding.

The test consists of four timed sections which are taken together in one sitting. Test takers move from task to task either by selecting the 'NEXT' button on completion of a task, or by being automatically moved to the next task at the end of the allotted time.

5.2.1 provides an overview of GATEWAY before each test section is looked at in detail. Every item type is described and illustrated with a sample item as well as sample answers.

5.2.1 Overview of test sections

The following table provides an overview of each test section and shows the relationship between task types and test items, as well as domain coverage.

Skill	Part	Task type	Activity, strategy or competence (CEFR)	Testing point	Target language use situation in academic settings	Skills assessed	Item traits scored
S P E A K I N G	Part 1	Personal introduction	Being interviewed	Not assessed	Talking about familiar topics	Not assessed	Not assessed
	Part 2	Answer interview questions	Being interviewed	Providing detailed information Expressing personal opinions	Discussions Immediate response	Listening and Speaking	Content Fluency Pronunciation Vocabulary Control Grammatical Accuracy
	Part 3	Describe image	Explaining data/ visuals in speech	Organizing and sustaining extended discourse	Interpreting graphic information Providing descriptive/ instructive or informative/ expository spoken discourse	Speaking	Content Pronunciation Vocabulary Control Grammatical Accuracy
	Part 4	Read aloud	Presenting written information	Recognizing academic vocabulary Identifying the writer's purpose and the tone	Understanding at a literal level	Reading and Speaking	Content: correct/incorrect Fluency Pronunciation

Skill	Part	Task type	Activity, strategy or competence (CEFR)	Testing point	Target language use situation in academic settings	Skills assessed	Item traits scored
R E A D I N G	Part 1	Multiple choice single answer	Reading for information and argument Identifying cues and inferring	Understanding and interpreting the gist and details of a text on an academic subject Analyzing the purpose of an academic text Evaluating or inferring meaning and tone	Understanding academic text Understanding at a literal level Critical reading Identifying assumptions or prejudices	Reading	*Dichotomously scored*
	Part 2	Collocation: Select the most appropriate word	Identifying (grammatical and lexical) cues Vocabulary range	Reading comprehension Collocational competence	Comprehending and producing academic writing	Reading	*Dichotomously scored*
	Part 3	Reorder paragraphs	Identifying (grammatical and lexical) cues Coherence	Structure (coherence and cohesion) and development Understanding organizational patterns Interpreting cohesive devices Reconstructing coherent discourse	Skimming Structuring discourse	Reading	*Dichotomously scored*

Skill	Part	Task type	Activity, strategy or competence (CEFR)	Testing point	Target language use situation in academic settings	Skills assessed	Item traits scored
	Part 4	Word formation: Write the correct form of the word	Identifying (grammatical) cues Accuracy	Understanding academic vocabulary Understanding grammatical structures Producing derived word forms	Academic writing	Reading and Writing	*Dichotomously scored*
L I S T E N I N G	Part 1	Identify opinions of participants	Understanding conversation between other people	Comprehending a discussion	Discussion Debates Persuasive/argumentative spoken discourse	Listening	*Dichotomously scored*
	Part 2	Express your opinion (speaking)	Sustained monologue: Putting a case	Extended oral production Expanding an argument with supporting detail	Short talks	Listening and Speaking	Content Fluency Pronunciation Vocabulary Control Grammatical Accuracy
	Part 3	Multiple choice single answer	Understanding broadcast audio and recordings Identifying cues and inferring	Comprehending academic spoken English of a variety of accents Identifying the purpose of an academic recording Inferring meaning	Following academic lectures, talks, or presentations	Listening	*Dichotomously scored*

Skill	Part	Task type	Activity, strategy or competence (CEFR)	Testing point	Target language use situation in academic settings	Skills assessed	Item traits scored
	Part 4	Present an aspect of a lecture	Understanding as a member of a live audience Relaying specific information	Comprehending spoken academic English Paraphrasing and summarizing key aspects in accurate spoken discourse	Relaying academic content from a lecture to fellow students	Listening and Speaking	Content Fluency Pronunciation
	Part 5	Debate: Take notes	Understanding as a member of a live audience Note-taking	Comprehending and processing spoken academic English Recording of key aspects in writing	Following academic debates, seminars, presentations and extracting key information by taking notes	Listening and Writing	Note-taking (lectures, seminars, meetings, etc.)
	Part 6	Debate: Express your opinion in writing	Reports and essays Propositional precision Coherence and cohesion	Presenting a coherent opinion and persuading the reader that a position is credible	Attending and participating in talks, debates, panel discussions Relaying/building on spoken academic content	Listening and Writing	Content Grammatical accuracy Vocabulary control
W R I T I N G	Part 1	Summarize written text	Reading for information and argument Mediating text (Processing a text in writing)	Comprehending and processing written academic English Identifying key aspects Paraphrasing and summarizing key aspects in accurate writing	Relaying spoken academic content	Reading and Writing	Content Grammatical accuracy Vocabulary control

Skill	Part	Task type	Activity, strategy or competence (CEFR)	Testing point	Target language use situation in academic settings	Skills assessed	Item traits scored
	Part 2	Write essay	Reports and essays Thematic development Propositional precision Coherence and cohesion	Overall written production Writing clear, detailed texts on an academic subject Synthesizing and evaluating relevant information and arguments Presenting a coherent argument and/or persuading the reader that a position is credible	Writing argumentative or persuasive essays	Writing	Content Grammatical accuracy Vocabulary control Vocabulary range Coherence and cohesion

Of the item types listed in the table above, the following are considered *integrated* item types that emulate the language learners encounter in educational settings. Students are expected to combine several language skills to participate successfully in their studies. Integrated item types are used to assess a combination of language skills.

Integrated item type used	Combined skills assessed
Read aloud	Reading & Speaking
Summarize written text	Reading & Writing
Write the correct form of the word	Reading & Writing
Take notes	Listening & Writing
Present an aspect of a lecture	Listening & Speaking
Answer interview questions	Listening & Speaking
Express your opinion	Listening & Speaking

5.2.2 Speaking Section

The Speaking section consists of items designed to test a wide range of general and academic speaking skills, including communicating about concrete and abstract topics, describing images, providing reasons and explanations and sustaining extended discourse using academic English.

Format	Test takers respond in real-time to a variety of general and academic prompts. Test takers are expected to respond spontaneously with limited or no preparation time. All test-taker responses are recorded and reviewed by proctors and assessed by human raters.
Length	25 minutes
Item types	The speaking section consists of five different item types assessing one language skill or a combination of two (integrated item types). The item types represent functions or situations test takers will experience in academic settings.
Sources	Images, texts and recordings are based on academic content of subjects in the humanities, social science and natural science.
Scoring	Item trait scoring using analytic 0 - 6 point scales based on the CEFR

5.2.2.1 Speaking Part 1 – Personal introduction

Item type and format	Test takers are asked two questions which allow them to introduce themselves and talk about a range of topics, such as work, studies, interests and future aspirations.
Time to answer	Test takers have 60 seconds per question to respond.
Task focus	This part of the test focuses on the test taker's ability to communicate about familiar and routine matters.
Function	This part of the test focuses on putting test takers at ease and allowing them to present themselves to the score users. The recording of the personal introduction is made available on the administrator dashboard for the university to review.
Item traits	N/A
Scoring	The responses are not scored.
Item count	2

5.2.2.2 Speaking Part 2 – Answer interview questions

Task Type and Format	The test taker will be asked intermediate to advanced questions over a range of less concrete/less familiar and more abstract/more complex topics which are academic in nature. Test takers will be expected to respond spontaneously.
Time to answer	Test takers have 60 seconds per question to respond.
Task focus	This part of the test focuses on the test taker's ability to communicate about more abstract topics in detail.
Function	Assesses the test taker's ability to • respond to questions • provide detailed information • express personal opinions on unfamiliar topics.

Item traits	- Content - Vocabulary Control - Grammatical Accuracy - Pronunciation - Fluency
Scoring	Item trait scoring using analytic 0-6 point scales based on the CEFR
Item count	1

Sample task: Answer interview questions

Rubric   You will hear three questions about education. You will have 60 seconds to answer each question.

Input

1. How well has education prepared you so far for the future?
2. What abilities or skills do you think students need to succeed at university?
3. To what extent do you think technological advances have a positive influence on education?

Sample answer

1. Several of my A-level classes have tied into real world experiences that help prepare for my planned studies at university or my future career in business. A recent example is a case study we explored in my A-level business class, where we reviewed the merger between a large national firm and a smaller regional start-up. It provided valuable insight into the financial models used for making the merger. This is something I will be able to apply in practice.
2. Nowadays, I believe that perseverance and resilience are most important to succeed at university. And of course, independence. We need to be able to cope with setbacks, learn from failure and move forward. It is not helpful to expect that we will always succeed. We need to do our best but not be discouraged when things don't work out as expected. Education is about learning and making mistakes is part of learning.
3. Yes, I do believe technology has mainly positive effects on education as it empowers the individual. Technology has been shown to increase student motivation, student engagement as well as student collaboration. In addition, technology also provides for more hands-on learning opportunities and allows for learning at all levels and thereby increasing students' confidence and technology skills.

5.2.2.3 Speaking Part 3 – Describe image

Describe image: Statistical data

Item type and format	Test takers are shown a graph, chart, map or table.
Time to answer	Test takers have 60 seconds per image to respond.
Task focus	This part of the test focuses on test takers' ability to describe in detail an image from an academic source, such as textbooks, journal articles, academic websites, etc. Test takers are asked to describe and comment on what they can see in the image. They may need to suggest a theme or trend, give opinions, or make predictions.
Function	Assesses test takers' ability to ● interpret graphic information ● provide descriptive or informative spoken discourse ● organize and sustain an extended discourse ● provide detailed information ● express thoughts on more abstract topics.
Item traits	● Content ● Vocabulary Control ● Grammatical Accuracy ● Pronunciation
Scoring	Item trait scoring using analytic 0–6-point scales based on the CEFR
Item count	1

Sample tasks: Describe image (Statistical data)

Level   B1

Rubric   Look at the image below. In 30 seconds, describe and comment on what the bar chart shows. You will have 60 seconds to give your response.

Input

Frequency of Internet Use in the UK, 2006-2020
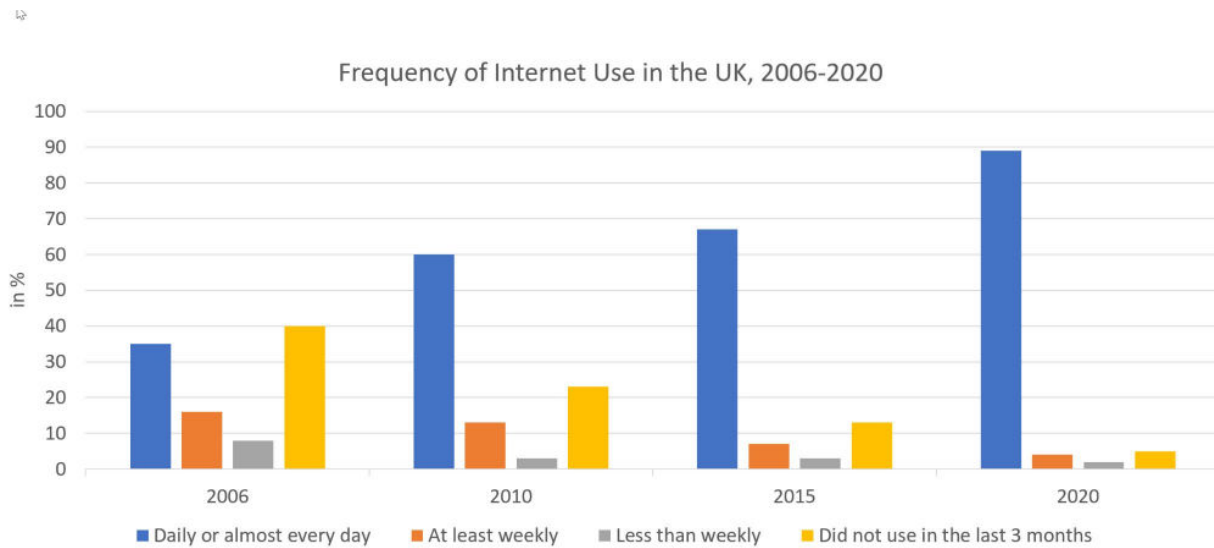
Source   UK Office for National Statistics ons.gov.uk

Sample answer

The bar chart shows the frequency of internet use in the UK from 2006 to 2020. Over this period daily use of the internet has almost tripled and, by 2020, 89 per cent of people were using the internet daily. The most dramatic increase in daily use of the internet happened between 2006 and 2010 when usage went from 35 to 60 percent of people. Correspondingly, from 2006 to 2020 patterns of less frequent use of the internet (once a week, less than once a week, use every few months) declined. This shows us that internet usage has really become a daily occurrence for the majority of the UK population. People are more dependent now on the internet as a source of information and many daily transactions we conduct such as shopping or contacting people happens over the internet.

Describe image: Process

Task type and format	Test takers are shown an image of a process. They have 30 seconds to look at the image followed by 60 seconds to describe and interpret it.
Time to answer	Test takers have 60 seconds per image to respond.
Task focus	This part of the test focuses on test takers' ability to describe in detail an image from an academic source, such as textbooks, journal articles, academic websites, etc. Test takers look at a visualization of a process. Test takers are asked to describe and interpret what they can see.
Function	Assesses test takers' ability to ● provide descriptive/instructive or informative/expository spoken discourse ● organize and sustain an extended discourse ● provide detailed information ● express thoughts on more abstract topics.
Item traits	● Content ● Vocabulary control ● Grammatical accuracy ● Pronunciation
Scoring	Item trait scoring using analytic 0–6-point scales based on the CEFR
Item count	1

Sample task: Describe image (Process)

Level   C1

Rubric   Look at the image of a process below. In 30 seconds, describe and comment on the process. You will have 60 seconds to give your response.

Input

Source   Inspired by
https://brookfieldinstitute.ca/wp-content/uploads/BrookfieldInstitute-PIP-Landscape-1.pdf
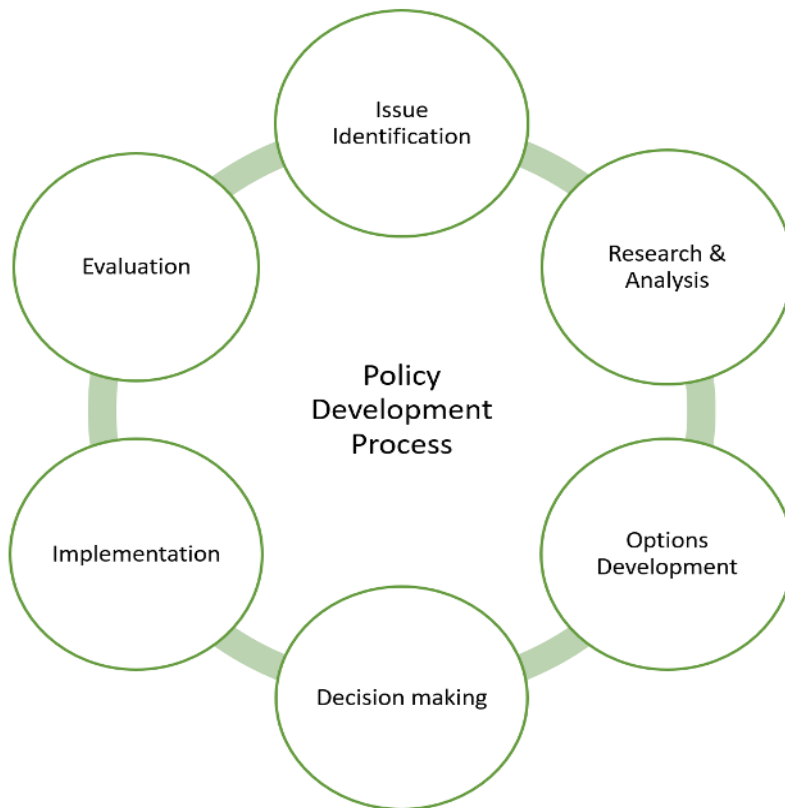
Sample answer

This image shows the policy cycle or how policies are developed. There are six stages to the cycle. When, for example, the government passes a law or takes an action, we refer to this as a policy, i.e., a consistent way of doing things. The first step in the process is to identify the issue followed by conducting research and analysis of the identified issue. The next step of the policy process is policy formulation, whereby potential policies are developed. Once a decision on a specific policy has been made, the process moves to policy implementation, which allows the policy to be adopted. The final phase of the policy development process is policy evaluation, when the successes or failures of a particular policy are looked at.

5.2.2.4 Speaking Part 4 – Read aloud (integrated)

Task type and format	Test takers read aloud an academic text of 50-70 words as clearly and fluently as possible.
Time to answer	Test takers have 60 seconds to read out each text.
Task focus	This part of the test focuses on test takers' ability to fluently read a text aloud using correct pronunciation and intonation.
Function	Assesses test takers' ability to ● read fluently ● recognize academic vocabulary; and ● identify the writer's purpose and tone of the text.
Item traits	● Content (count of replacements, omissions and insertions of words) ● Pronunciation ● Fluency
Scoring	Item trait scoring using analytic 0–6-point scales based on the CEFR
Task count	1

Sample tasks: Read aloud

Level   C1

Rubric   Read the text aloud as clearly and fluently as possible.

Input

> Sole traders do not have a separate legal existence from the business. In the eyes of the law, the business and the owner are the same. As a result, the owner is personally liable for the firm's debts and may have to pay for losses made by the business out of their own pocket. This is called unlimited liability.

Source   https://www.bbc.co.uk/bitesize/guides/z4br87h/revision/1

5.2.3 Reading Section

The Reading section consists of items designed to test a wide range of academic reading skills, including reading for main ideas, reading for detail, skimming, understanding logical argument, and recognizing writers' opinions, attitudes, purpose, and understanding academic vocabulary.

This section only includes independent reading skills items.

Format	Test takers engage with a variety of academic texts covering a range of subjects and answer questions presented in multiple formats. All test-taker responses are assessed automatically.
Length	Approx. 40 minutes
Item types	The reading section consists of five different item types assessing one language skill or a combination of two (integrated item types). The integrated reading-into-writing and reading-into-speaking item types are, however, part of the writing and speaking test sections respectively. The item types represent functions or situations test takers are likely to experience in academic settings.
Sources	Texts are taken from academic books, journals or magazines. All texts deal with issues which are recognizably appropriate and accessible to test takers entering undergraduate or postgraduate courses. All the information needed to answer the items are contained in the text. The passages may be written in a variety of styles. For example, narrative, descriptive, or argumentative.
Scoring	The item types in this section are dichotomously scored: Correct or incorrect: Test takers will receive the maximum, i.e., 1 score point for each item. If the response is incorrect, they will receive no score points.

5.2.3.1   Reading Part 1 – Multiple choice single answer

Task type and format	After reading a passage of 350-450 words, test takers answer four multiple-choice questions by selecting the correct option out of a possible four.
Task focus	This section focuses on reading comprehension and the ability to analyze, interpret and evaluate the content of an academic text.
Functions	Assesses test takers' ability to understand and interpret the content of a text on an academic subject. Items focus on testing the understanding of: ● specific information ● gist ● main ideas or themes ● opinions or attitudes ● purpose ● inference
Scoring	Correct/incorrect 1 Each correct response 0 Minimum score
Task count	2

Sample task: Multiple choice single answer

Level   C1

Rubric   Read the text. Choose the best answer for each question. Only one option is
             correct.

Input

> ### Francis Bacon
>
> Francis Bacon was an English lawyer, statesman, essayist, historian, philosopher who lived during the 16th and 17th centuries. He is often referred to as a "universal genius" which is an accolade reserved for those who make significant, original contributions to more than one professional discipline or areas of learning. And judging by his accomplishments he merits this designation. In our own era Bacon would be acclaimed as a "public intellectual," though his personal record of service and authorship would

certainly dwarf the achievements of most academic and political leaders today.

Relatively early in his career Bacon judged that, mainly because of a deep respect for the past (as well as to an excessive absorption in cultural vanities and frivolities), the intellectual life of Europe had reached a kind of standstill. Yet he believed there was a way beyond this stagnation if persons of learning, equipped with new methods and insights, would simply open their eyes and minds to the world around them. This was the basic argument of his 1605 book The Proficience and Advancement of Learning, arguably the first important philosophical work to be published in English.

A key insight of Bacon's in The Proficience and Advancement of Learning is that knowledge is power, and when embodied in the form of new technical inventions and mechanical discoveries it is the force that drives history. Though it is hard to pinpoint the birth of an idea, essentially, the modern understanding of technological "progress" (in the sense of a steady, cumulative, historical advance in applied scientific knowledge) can be traced back to this work. And it is all the more remarkable that it was conceived when most English and European intellectuals were deploring what they saw around them as the numerous signs of modern degradation and decline.

Bacon wrote literary as well as scientific and philosophical works. Despite the fanatical claims of a few admirers, he did not author the works traditionally attributed to William Shakespeare. However, he made important contributions to English literature and influenced the development of English prose style. Interestingly his style achieves its air of ease and clarity more through its balanced cadences, natural metaphors, and carefully arranged symmetries than through the use of plain words, commonplace ideas, and simple syntax. Furthermore, just as Bacon's individual style and living habits were prone to extravagance and never particularly austere, so in his literary works he was never quite able to resist the occasional grand word.

Word count = 408

Source   https://iep.utm.edu/bacon/ amended in line with vocabulary analysis

Items

1. The author thinks that Francis Bacon's reputation as a universal genius

a. should be reconsidered.
b. may be slightly exaggerated.
c. is entirely well-deserved.
d. has strengthened over time.

> 2. According to the author, Bacon believed that people learning in the 17th century
>
> a. lacked the equipment necessary to study.
> b. were afraid of adopting new ways of thinking.
> c. gave insufficient attention to historical events.
> d. needed to be more curious about their surroundings.
>
> 3. In paragraph 3, the author argues that *The Advancement of Learning*
>
> a. established a new concept.
> b. is still relevant in current times.
> c. led to a number of inventions.
> d. was misunderstood by scholars.
>
> 4. What does the author say about Bacon's literary style?
>
> a. It uses straightforward language.
> b. It reflects a part of his personality.
> c. Its complexity is largely unappreciated.
> d. It is influenced by Shakespeare's work.

Key   1. c   2. d   3. a   4. b

5.2.3.2 Reading Part 2 – Select the most appropriate word

Task type and format	Test takers complete a gapped reading text of up to 250 words by choosing a single correct answer for each of the six gaps from a set of four options.
Task focus	This section tests collocational competence in academic English. It focuses on test takers' ability to use context as well as collocational knowledge to complete a gapped text.
Functions	Assesses test taker's ability to ● understand academic vocabulary ● identify collocating words; and ● infer meaning of unfamiliar words.
Scoring	1 Each correct response 0 Minimum score

Item count	6

Sample task: Select the most appropriate word

Level   C1

Rubric   Read the text. Complete each gap by selecting the most appropriate word.

Input

> Animal Migration
>
> Animal migration is a pattern of behavior in which animals, often in groups, travel from one habitat that they are accustomed to living in to another in search of food or better conditions, or to address reproductive needs. It is a phenomenon far grander and more patterned than an animal's general wandering. It usually represents purposeful and directional (1) _____ travel with long-deferred rewards.
>
> Although migrating behavior seems to involve planning and intention, it all comes down to (2) _____ instinct. Not all animals migrate, but those that do include species of birds, mammals, reptiles, amphibians, fish, insects, and crustaceans. Migrating species differ in their reason for migrating, their types of migration, the distances traveled, and their navigation skills.
>
> Biologists have identified five characteristics that apply, in (3) _____ degrees and combinations, to all migrations. Migrations are prolonged movements that carry animals outside (4) _____ habitats; they tend to be linear, not zigzagged; they have a defined beginning and end; they involve special behaviors of preparation, such as overfeeding in the run-up; and they demand special energy reserves.
>
> Migrating animals need to devote their (5) _____ attention to the greater mission. In order for these animals to survive and complete the migration successfully, it is crucial that they are not distracted by temptations, and that they remain undeterred by challenges that would discourage virtually (6) _____ other animals in order to survive and complete the migration successfully.

Source   Adapted from https://www.nationalgeographic.com/magazine/article/great-migrations

Options for items 1–6

1.	A	unlimited	B	domestic	C	collective	D	corporate
2.	A	maternal	B	natural	C	paternal	D	protective
3.	A	varying	B	unequal	C	greater	D	minor
4.	A	suitable	B	critical	C	preferred	D	familiar
5.	A	undivided	B	close	C	urgent	D	personal
6.	A	few	B	some	C	most	D	all

Key   1. C 2. B   3. A   4. D   5. A   6. D

5.2.3.3 Reading Part 3 – Write the correct form of the word

Task type and format	Test takers complete six gaps in a reading text of 150 to 200 words with the correct form of each word. The items are set within a continuous text, providing contextual support and allowing for testing negative forms and thus testing overall understanding of the sentence or paragraph.
Task focus	This section focuses on test taker's ability to use context and grammatical knowledge to correctly form the word that completes each associated gap.
Functions	Assesses test takers' ability to ● understand academic vocabulary ● identify collocating words; and ● infer meaning of unfamiliar words.
Scoring	1 Each correct response 0 Minimum score
Item count	6

Sample task: Write the correct form of the word

Rubric   Read the text. Complete each sentence using the correct form of the word given in brackets.

Input

> Hippolyte Blancard, a pharmacist and amateur photographer, documented Parisian architecture leading up to the 1889 World's Fair, an international event (1) (showcase) _____ new innovations, geographic and scientific discoveries, and works of art. This series of photographs depicts the construction of the Eiffel Tower, which was conceived as the (2) (enter) _____ to the World's Fair.
>
> Blancard's photographic documents cover the tower's progression, from July 1887 through April 1889. There was a strong reaction to the Eiffel Tower among the general public, some of whom thought it was (3) (sight) _____ and a "stain" on the Paris cityscape. One critic called it "a truly tragic street lamp." Many artists, however, (4) (embrace) _____ the Eiffel Tower as a symbol of modernity and the avant-garde.
>
> Gustave Eiffel, the tower's builder and designer, responded to the criticism in a newspaper interview, saying, "For my part I believe that the tower will possess its own beauty. Are we to believe that because one is an engineer, one is not preoccupied by beauty in one's constructions or that one does not seek to create (5) (elegant) _____ as well as solidity and durability?" Initially, the plan was to demolish the tower, but it gradually became a (6) (define) _____ icon of the Paris cityscape.

Source   Adapted from: The Museum of Modern Art, New York, USA.
          https://www.moma.org/learn/moma_learning/ hippolyte-blancard-untitled-construction-of-the-eiffel-tower-april-1889/

Key   (1) showcasing (2) entrance (3) unsightly (4) embraced (5) elegance (6) defining

5.2.3.4   Reading Part 4 – Reorder paragraphs

Task type and format	Test takers reconstruct the original order of a reading text of up to 250 words by dragging and dropping text boxes in a single correct order.
Task focus	This section focuses on the test taker's ability to understand the organization, coherence and cohesion of an academic text.
Functions	Assesses the test taker's ability to ● understand organizational patterns ● interpret cohesive devices ● reconstruct coherent discourse.
Scoring	1 Each correct response 0 Minimum score
Item count	2

Sample task: Reorder paragraphs

Rubric   The paragraphs of an academic text have been placed into a random order. Sort the paragraphs into a coherent order by dragging each text box into its correct place.

Input

> (1) This amounted to a great achievement hardly to be expected in his day and age, although nowadays Galileo Galilei's conclusion may appear to some people to be trivial and insignificant. Not long after his lunar observations, the skies which had previously been elusive revealed more of their extraordinary mysteries.
>
> (2) Three hundred and fifty years before the first men looked down on the surface of the moon from close quarters, Galileo Galilei's newly built telescope enabled him to look at the edge of the hitherto mysterious sphere.
>
> (3) Galileo observed the striking planet of Jupiter and nestling next to it, he saw four little points of light circling the distant planet. Our moon, it appeared, perhaps unfortunately in the eyes of those fearful of what the discovery might mean, was not alone.

(4) He saw that the apparently lifeless surface of our moon was not smooth and round as assumed, but actually bumpy and imperfect. He realized that although the moon might appear inactive, resembling a still life painted by the hand of a cosmic artist, it was a real world, perhaps not even so very different from our own.

Source https://solarsystem.nasa.gov/news/307/galileos-observations-of-the-moon-jupiter-venus-and-the-sun/

Key   2, 4, 1, 3

5.2.4 Listening Section

The listening section consists of items designed to test a wide range of academic listening skills. These include listening for gist, main ideas, and detail, understanding logical argument, recognizing the speaker's opinions, attitudes and purpose, and understanding academic vocabulary. Recordings are based on or inspired by academic lectures, presentations and discussions, etc.

Format	Test takers are presented with audio or video recordings of simulated academic lectures, presentations and conversations, etc. and will answer questions presented in multiple formats.
Item types	The listening section consists of six different item types. Some item types are integrated and assess both listening as well as writing skills or speaking skills. The item types represent functions or situations test takers experience in academic settings.
Length	Approx. 30 minutes
Sources	Audio and video recordings taken from simulated conversations, lectures, academic podcasts and discussions, news casts, and announcements.
Scoring	The item types in this section include two types of scoring: 1. Correct or incorrect: Test takers will receive the maximum of 1 score point for each item type. If the response is incorrect, they will receive no score points. 2. Partial credit: If the response is correct, the test taker will receive the maximum score points available for each item type. If the response is

	partially correct, they will receive partial credit. If the response is incorrect, the test taker will receive no score points.

5.2.4.1 Listening Part 1 – Identify opinions of participants

Task type and format	While listening to an audio recording, test takers answer six multiple-choice matching questions on the content or purpose of the recording by selecting one response per item.
Task focus	This section focuses on the ability to comprehend the content or purpose of an academic conversation, discussion or debate by identifying the participants' opinions correctly.
Functions	Assesses the test taker's ability to understand ● persuasive/argumentative spoken discourse ● agreement/disagreement ● opinions.
Scoring	Correct/incorrect 1 Each correct response 0 Minimum score
Task count	1

Sample task: Identify opinions of participants

Rubric   You will hear a conversation between two speakers. For statements A-F, decide whether the opinions are expressed by only one of the speakers or both. Click the appropriate button for each question. Only one option is correct.

Input   Audio

Transcript (scripted)

O	Susan, do you remember when we first arrived on campus?
S	O yes, that was a rollercoaster journey, wasn't it?
O	The initial chaos of rooms being wrongly assigned. I still can't quite believe how that could have happened.
S	And the wi-fi wouldn't work for days on end. Then they lost my registration form for the lab work and I almost lost my place.
O	Fortunately, the first three weeks were an exception. It all calmed down pretty quickly.

S	So true! We settled into our courses quite quickly. From the very beginning classes focused on theory and practice. It's so important when you study biology to actually do field work right from the start.
O	Yes, lucky you - no practical courses for me though. Philosophy doesn't really go beyond theory at uni.
S	You are always complaining that your student life probably looks the same as it did when Thoreau studied at Harvard - books, essays, debates, libraries, books, essays, debates …
O	Ha ha, at least I've joined lots of clubs. This is how I made friends initially and built confidence.
S	As if you needed any! Honestly though, we didn't just focus on your classes, you are right.
O	And, you know, once you build up the confidence in clubs, it'll be easier to speak up in class or present your work in tutor groups - which are quite large. Too large for my liking.
S	Yes, mine too. Their size should be limited to allow for more efficient learning…
O	And debating - in my case. What will my job prospects be if I can't argue effectively!
S	Limited at best! [laughs]
O	Thanks Susan. It's easy for you to say…
S	I know. The practical placement at the end of my course is a great opportunity. I'm sure employers will come knocking on my door. I should start thinking about where to apply.
O	Common now--hold your horses - we are only in our first year after all!

Items A - F

Susan and Oliver are talking about their undergraduate studies.

		Susan	Oliver	Both
A	The first few weeks were chaotic.			
B	The course offers theoretical and practical knowledge.			
C	In some respects, student life is similar to what it was in the past.			
D	Extracurricular activities build confidence.			
E	Students would benefit from smaller tutor			

	groups.				
F	The course structure improves employability.				

Key   A - both   B - Susan   C - Oliver   D - Oliver   E - both   F - Susan

This integrated item type is followed by a speaking task.

5.2.4.2   Listening Part 2 – Express your opinion (integrated)

Item type and format	Test takers are asked to share their opinions about ideas, issues, and themes raised in the corresponding listening comprehension task (Part 1 Identify opinions of participants). Test takers respond spontaneously in real-time.
Time to respond	Test takers have 90 seconds to give their response.
Task focus	This section focuses on the test taker's ability to give reason and support for their position.
Function	Assesses the test taker's ability to ● state opinion ● give reason ● expand an argument with supporting detail.
Item traits	● Content ● Vocabulary Control ● Grammatical Accuracy ● Pronunciation ● Fluency
Scoring	Item trait scoring using analytic 0 - 6-point scales based on the CEFR
Item count	1

Sample tasks: Express your opinion

Level   B2

Rubric   You have just listened to a conversation about two students' undergraduate courses. Now express your opinion on the following topic. You have 90 seconds for your response.

Input

> University courses should ensure that students are well-prepared for their chosen career. Do you agree or disagree with this statement? In your response, address the opinions expressed on this topic in the conversation between Susan and Oliver, and support your point of view. You may include examples from your own experience.

Sample answer

I disagree with this statement and argue that it depends on the subject as illustrated in the conversation between Susan and Oliver. I don't think that the objective of all university courses should be to ensure students are well-prepared for their future career. First and foremost, university courses should provide the theoretical foundation that is required in understanding a particular subject. Ideally, this would lead to practical applications of the learned knowledge. Without theoretical understanding, however, progress in the field may be limited. Once students grasp the theoretical framework, they are often able to apply their knowledge to real life situations. This means university education should equip students with knowledge that can be applied across a variety of contexts rather than prepare students for one specific line of work.

5.2.4.3 Listening Part 2 – Multiple choice single answer

Task type and format	While listening to an audio recording, test takers answer four multiple-choice questions on the content or purpose of the recording by selecting the correct option out of a possible four.
Task focus	This section focuses on the ability to comprehend the content or purpose of an academic audio recording in detail.
Functions	Assesses the test taker's ability to • comprehend academic spoken English of a variety of accents • identify the purpose of an academic recording • infer meaning
Scoring	Correct/incorrect 1 Each correct response 0 Minimum score
Item count	4

Sample task: Listening Multiple choice single answer

Level   C1

Rubric   Listen to the recording and choose the best answer for each question. Only one option is correct.

Input

Ants are some of the most successful insects on the planet, and it is believed this is because of the division of labor and the way certain ants specialize in certain tasks within a colony.

Today's broadcast is on *Temnothorax albipennis* or the rock ant. They build nests in rock crevices, hence their name. Rock ants are small, light brown in color and are found, for example, in coastal areas of the United Kingdom.

Today we want to focus on two lesser-known facts about the species brought to light by recent studies. Firstly, the social learning process referred to as tandem running and secondly, the phenomenon of behavioral lateralization.

Rock ants have been observed teaching each other through a process known as tandem running. An experienced forager—that's an ant that goes out in search of food—leads a less experienced nest-mate to a newly discovered resource such as food or an empty nest site. The follower obtains knowledge of the route by following in the footsteps of the tutor, maintaining contact with the tutor's antennae. Both leader and follower are aware of the progress made by the other. So, the leader will slow down when the follower lags behind and speed up when the follower gets too close.

Rock crevices, potential homes for rock ants, are plentiful. When rock ants explore a new nest site, and that is tonight's second lesser-known fact, they show a leftward turning bias. Research from the University of Bristol in the UK has revealed how ants engage in the phenomenon known as behavioral lateralization, which refers to a species' preference to use a certain side of the body for certain tasks.

Edmund Hunt, along with his colleagues, studied how rock ants find their way as they explore rock cavities. The researchers found that rock ants were significantly more likely to turn left than right when exploring new spaces. Such behavioral lateralization may also relate to physical asymmetries - such as the size of eyes or the length of legs. The researchers found that ants turning left tended to have better vision in their right eye. They prefer to walk with their inferior eye pointing toward the wall, so when they come to a branch, they follow the wall along to the left.

Hunt also pointed out that as their world is maze-like, consistently turning one way is a great strategy to search and exit any maze. As their nest-mates prefer turning left, too, this approach should also create safety in numbers helping the ants to safeguard nest-mates from predators during house hunting.

So, all in all, always turning instinctively left is a very smart evolutionary adaptation indeed.

word count: 444

Source   Adapted from
https://www.brightsurf.com/news/article/041118453892/having-one-eye-better-than-the-other-may-explain-ants-left-bias.html
https://en.wikipedia.org/wiki/Temnothorax_albipennis

Items

1. Which statement about the species' nest sites is correct?
    A. They are usually simple and small.
    B. They are safer than other ants' nests.
    C. There are not many suitable locations for them.
    D. They are found in cracks that form in rocks.

2. What is the main purpose of tandem running?

    A. It enables ants to work faster.
    B. It helps keep young ants safe.
    C. It prevents ants from starving.
    D. It allows for information transfer.

3. According to Edmund Hunt, behavioral lateralization may help the ants

    A. build their nests.
    B. find their way.
    C. locate food sources.
    D. communicate with other colonies.

4. What is the intent of this broadcast?

    A. To present research findings
    B. To show how rock ants detect predators
    C. To explain how rock ants build nests
    D. To disprove the claims of earlier research

Key   1D   2D   3B   4A

5.2.4.4 Listening Part 4 – Present an aspect of a lecture

Task type and format	Test takers listen to an excerpt of 60–75 seconds taken from an academic lecture before being asked to relay specific information from the lecture. Test takers will have 60 seconds to prepare their response and then 60 seconds to respond.
Task focus	This part of the test focuses on test takers' ability to follow an academic lecture and present one specific aspect.
Function	Assesses test takers' ability to ● understand academic vocabulary ● identify relevant information ● present specific information clearly.
Item traits	● Content based on the CEFR scales *Understanding as a member of a live audience* and *Overall oral production* ● Pronunciation ● Fluency
Scoring	Item trait scoring using analytic 0 - 6-point scales based on the CEFR
Task count	1

Sample task: Present an aspect of a lecture

Rubric   You are going to hear two short excerpts from two academic lectures. You will be asked to present one specific aspect of each lecture. After listening to the lecture, you will have 60 seconds to prepare your spoken response and then 60 seconds to record your response. You may take notes in the area provided by the test. Do not use pen or paper to take notes.

Input   Audio

Transcript

> The subject of today's lecture is Culture Shock.
>
> Culture shock, as you know, is the term used to describe the experience many people have when they travel to another country, and it can be seen as a manifestation of group pressure in action.

It is a good example of group pressure, because it shows what happens when an individual suddenly experiences different cultural rules — the rules of another cultural group. Now culture shock is a complex phenomenon, but I'm going to focus on three main ideas in this lecture.

First of all, we will consider the reasons why people experience culture shock. Second, I will describe the different stages of this experience. Finally, I'll mention some possible applications of this research because although you might think that culture shock affects, say, only travelers, that is not the case. In fact, cross-cultural studies have immense practical value for modern society.

First, then, why do people experience culture shock? *[fading out]*

Source https://www.slideshare.net/Boojie/academic-encounters-lecture-scripts (excerpt)

Sample answer (transcript)

The lecturer talks about culture shock as a form of group pressure in action. According to the lecturer, culture shock can be considered group pressure because an individual from a different culture has to conform to different cultural rules — the rules of another cultural group. This can be experienced as pressure.

5.2.4.5 Listening Part 5 – Take notes (integrated)

Task type and format	Test takers watch a short video of a group discussion or debate on an academic topic. Test takers are asked to take notes on whether the participants agree or disagree with the topic and the reasons they provide to support their opinion.
Task focus	This part of the test focuses on test takers' ability to extract key information from an academic recording and note them down.
Function	Assesses test takers' ability to ● comprehend and process spoken academic English of various accents; ● identify key aspects made by a variety of discussants in a debate; ● make notes in writing.
Item traits	● Content based on the CEFR scale *Notetaking (lectures, seminars, meetings, etc.)* ● Grammatical accuracy

	● Pronunciation ● Fluency
Scoring	Item trait scoring using analytic 0 - 6-point scales based on the CEFR
Task count	1

Sample task: Take notes

Level B2

Rubric You are about to hear a debate led by a moderator on whether the increasing use of technology in the classroom is beneficial or not. Three students express their views. While listening to the recording, take notes on whether the three students agree or disagree with the topic and the reasons each of them provide.

Input video

Transcript

Moderator: Our topic for discussion is - the use of technology in the classroom has become more and more common. Do you think this is a good thing? Why or why not? Let's start with you, Tanisha.

Tanisha: I have to say I think it's a double edged sword. I think in some instances it's really good with kids who can handle it or even students, you know, college students who can handle it, but certain students can't. And I think that is where it lies. You just have to kind of base it from student to student to see if that's a good thing for them or not.

Moderator: Now that makes sense. What about you Carla? What do you think?

Carla: I would say it's more so a negative thing. It really kind of takes away from like the values of English and writing and things like that. You can type it down and stuff, but it really takes away like the enjoyment of writing like pen on paper kind of a thing and it kind of limits students I feel like. Now they only know how to use the technology and they don't know how to use any other form of writing or things like that.

Moderator: Well, that's an interesting perspective. What do you think, Dalton?

Dalton: Well, I think it's a good thing, but it gives and it takes. Like it matters if people even have access or if they have really the money to have their own computers or laptops at their house. And if they don't have that stuff and they're kind of struggling

on their own end. But overall, I think it's a good thing and we should transition toward it eventually.

Source   in-house recording

Sample answer

Tanisha	sees both positive and negative aspects varies from student to student
Clara	disagrees takes away from the enjoyment of other forms of learning/writing
Dalton	Mainly agrees But points out accessibility issues

5.2.4.6   Listening Part 6 – Express your opinion in writing (integrated)

Task Type and Format	After listening to a short debate on a specific topic and taking notes, test takers express their own stance on the topic in writing.
Task focus	This part of the test focuses on test takers' ability to express their opinion on an academic topic in writing of 120-150 words.
Function	Assesses test takers' ability to ● presenting a coherent opinion in accurate writing.
Item traits	● Content based on the CEFR scales *Reports and essays* and *Propositional precision* ● Vocabulary control ● Grammatical accuracy
Scoring	Item trait scoring using analytic 0 - 6 point scales based on the CEFR
Item count	1

Sample task: Express your opinion in writing

Level   B2

Rubric   You have just listened to a debate on whether the increasing use of technology in the classroom is beneficial or not. Now express your stance on this topic in writing. Your response should be between 120 and 150 words in length. You have 10 minutes to complete the task.

Input   The task links back to the video recording of the discussion presented in the previous task (5.2.4.5 Listening Part 5 - Take notes).

Sample answer

As technology changes the world, it is only to be expected that technology also changes the ways we teach our students. Making technology an integral part of the classroom benefits and supports learning and teaching practices.

One significant benefit of integrating technology in course design is that it creates a more engaging learning environment. Technology can help teachers to present topics in more interactive and creative ways. It helps with students' motivation and allows students to become more familiar with devices and software that they will use in their future careers. Technology also enables new channels of collaboration amongst students leading to new ways of interacting and learning together. In this way, technology also makes learning more accessible to students.

5.2.5   Writing Section

The writing section consists of item types designed to test a range of academic writing skills, including writing for a purpose, summarizing information and making inferences and predictions.

The test taker will be asked to write an essay in response to a question, point of view, argument or problem. Responses to this task should be in a standard essay format consisting of an introduction, body and conclusion. Topics are of general interest to, and suitable for, test takers entering undergraduate and graduate studies.

Please note that the integrated item type *Express your opinion in writing* is taken as part of the Listening section following the debate during which test takers are asked to take notes (cf. 5.2.4.6).

Format	Test takers are required to write responses in the formal register of academic English using correct grammar as well as appropriate vocabulary.
Item types	The writing section consists of three different item types of which two are integrated and assess both reading and writing skills. The item types represent functions or situations test takers will experience in academic settings.
Timing	45 minutes
Item count	3

Sources	Academic articles, articles providing political analysis, editorial, literary review, reports or extract from scientific books, and other similar online or print sources.
Scoring	Item trait scoring using analytic 0–6-point scales based on the CEFR

5.2.5.1   Writing Part 1 – Summarize written text (integrated)

Task Type and Format	After reading an academic text of 150-200 words, test takers write a brief summary of the main aspects.
Task focus	This part of the test focuses on the test taker's ability to comprehend information from an academic text and summarize the key aspects in writing of up to 80 words.
Function	Assesses the test taker's ability to ● comprehend and process written academic English ● identify key aspects ● paraphrase and summarize key aspects in accurate writing.
Item traits	● Content based on the CEFR scales *Processing text in writing* and *Reading for information and argument* ● Vocabulary Control ● Grammatical Accuracy
Scoring	Item trait scoring using analytic 0–6-point scales based on the CEFR
Item count	1

Sample task: Summarize written text

Level   C1

Rubric   Read the following text and summarize its main aspects in 50-80 words. You have 10 minutes to complete this task.

Input

> A group of countries often join together in a trade bloc to promote trade. This might be through relaxing barriers to trade, or even having a common currency, and increasing taxes on products brought in from outside the bloc, protecting them from outside competition. Examples of trading blocs include the EU, NAFTA and ASEAN. The European Union (EU) is currently made up of

> 27 member countries, has a population of over 500 million people and accounts for 25% of global GDP. It aims to promote trade by creating a single market where goods, money and people can travel freely between member states. It is believed that all member states will benefit through increased job creation and income. 19 of the member states also agreed to use a single currency, the euro. These countries make up an area called the Eurozone. Representatives of low-income countries, also known as developing countries, complain that they cannot get fair access to the European market, due to increased taxes and quotas on their products making them more expensive and therefore less marketable.

Source   https://www.bbc.co.uk/bitesize/guides/z8436fr/revision/6 (edited)

Sample answer

One notable trading bloc is the European Union. This trading bloc is made up of 27 member states. It was formed to promote trade by allowing the free movement of goods, money and people between member states to the detriment of developing countries that cannot gain fair access to the market.

5.2.5.2   Writing Part 3 – Write essay

Task Type and Format	Test takers write a persuasive or argumentative essay of 250-350 words in response to a stimulus or question. They have 25 minutes to complete the task. Responses to this task should be in a standard essay format consisting of an introduction, body, and conclusion. Topics are of general interest to, and suitable for, test takers entering undergraduate and graduate studies.
Task focus	This part of the test focuses on the test taker's ability to produce a clear, well-structured essay that develops an argument systematically with appropriate highlighting of significant points and relevant supporting detail.
Function	Assesses the test taker's ability to ● write clear, detailed texts on an academic subject ● synthesize and evaluate relevant information and arguments ● present a coherent argument and/or persuade the reader that a position is credible.
Item traits	● Content based on the CEFR scales *Reports & Essays* ● Thematic development & Coherence ● Vocabulary Range

	● Vocabulary Control ● Grammatical Accuracy
Scoring	Item trait scoring using analytic 0-6 point scales based on the CEFR
Item count	1

Sample task: Write essay

Level   C1

Rubric   Write an essay on the topic below. You have 25 minutes. Your response should be between 250 and 350 words in length.

Input

> Social media is responsible for widespread disinformation, which refers to the spread of inaccurate information with the intent to deceive. Discuss the extent to which you agree or disagree with this statement. Support your point of view with reasons and/or examples from your own experience.
>
> You may consider the following points:
>
> - Sources of disinformation
> - Who engages with disinformation
> - Why do we engage with disinformation

Sample answer

Social media has had a dramatic impact on modern life, on the way we interact with others and where we get our information from. Social media platforms have connected us in new and often meaningful ways.

However, stories and opinions can gain exposure with unprecedented speed, giving individuals around the globe continuous access to a near real-time conversation about both important and trivial matters. With the increasing popularity of a wide variety of internet-enabled devices and fast broadband, more and more people are now taking to social media on a daily basis. Indeed, many get their news through posts made on social networking sites.

Unfortunately, as there is little control on what is put on social media platforms, they have become a prevalent source of inaccurate news content. Content which may misinform us at best of times, or deliberately sets out to disinform us rather than provide us with facts and analysis needed for informed decision-making. Readers are often manipulated for political or economic reasons by fake news.

So far, society as a whole has had little experience in combating the issue that is currently spreading like wildfire. A recent example is the pandemic. In early 2020, a number of myths began making headlines regarding the supposed treatment of COVID-19. Many bogus claims were presented as fact, including the idea that eating more red meat or coconut oil could fight the virus.

Debunking such claims requires careful research, fact-checking, and media forensics. As not everyone has got the time or skills to do so, readers should rely on traditional news outlets that are rooted in investigative journalism.

In addition, it's important to understand the impact of sharing potential misinformation as it has a detrimental impact on public discourse by misleading and manipulating readers — whether intentionally or not. In conclusion, it is no longer a contentious topic that social media platforms have to be held responsible for their content if they want to meet their original purpose of bringing people together.

<div align="right">words: 330</div>

6 Test marking and scoring

The majority of test taker responses in the Reading and Listening section of GATEWAY are dichotomously scored by computer. All other responses are sent electronically to trained and certified English3 raters who rate them according to analytical criteria aligned to the level descriptors of the Common European Framework of Reference for Languages (CEFR). The ability of the test taker is estimated according to the responses given in relation to the difficulty of the item presented. The final ability estimate is derived once the complete set of test items has been delivered. Ability estimates are then converted to a standardized score and this is also reported in terms of a CEFR level.

Final test scores, including the overall score and four skill scores, i.e., Speaking, Reading, Listening and Writing, are reported on the English3 proficiency scale. Each of the four skill sections accounts for 25% of the overall score.

Analytic scoring is used for all extended responses to assess the speaking and writing proficiency of each test taker, which also includes the integrated item types of listening-into-speaking, listening-into-writing, reading-into-speaking and reading-into-writing. A separate score is awarded for each of the item traits (item-trait scoring), as opposed to assigning one holistic score for the task.

This allows for more exact reporting of proficiency, especially where language skills may be developing at different rates. Item-trait scoring also leads to greater reliability as each test taker is awarded a number of scores. Human raters are required to focus on each trait individually, thus ensuring that they are all addressing the same features of the test taker's performance.

Each rating scale for each item trait that is assessed ranks the performance along an ordinal scale of 0 to 6 and describes language proficiency as a series of levels against which the test taker's performance is judged.

Each level is characterized in terms of what the test taker can do with the language (functions which can be performed, for example, putting a case, presenting data, inferring, etc.) and their control of linguistic features (e.g., vocabulary, pronunciation, fluency, etc.). The development of the scoring rubrics is informed by the relevant descriptor scales of the CEFR, which ensures GATEWAY is closely aligned to the framework.

6.1 Overview of types of scoring

Item responses in the GATEWAY test are either computer-scored or scored by human raters. Scores for some item types are based on correctness alone. These are the dichotomously scored item types (correct/incorrect). Open responses are scored on individual item traits which assess qualitative and functional aspects of the response.

The table below shows which type of scoring is applied to each item type.

Com- ponent	Part	Task type	Skills assessed	Type of scoring	Item traits scored
S P E A K I N G	Part 1	Personal introduction	Speaking	Unscored	Not assessed
	Part 2	Answer interview questions	Speaking	Analytic scoring of individual item traits	Content Fluency Pronunciation Vocabulary Control Grammatical Accuracy
	Part 3	Describe image	Speaking	Analytic scoring of individual item traits	Content Pronunciation Vocabulary Control Grammatical Accuracy
	Part 4	Read aloud	Speaking	Analytic scoring of individual item traits	Content Fluency Pronunciation
R E A D I N G	Part 1	Multiple choice single answer	Reading	Dichotomously scored: correct/incorrect	
	Part 2	Collocation: Select the most appropriate word	Reading	Dichotomously scored: correct/incorrect	
	Part 3	Reorder paragraphs	Reading	Dichotomously scored: correct/incorrect	
	Part 4	Word formation: Write the correct form of the word	Reading	Dichotomously scored: correct/incorrect	

Component	Part	Task type	Skills assessed	Type of scoring	Item traits scored
LISTENING	Part 1	Identify opinions of participants	Listening & Speaking	Dichotomously scored: correct/incorrect	
	Part 2	Express your opinion	Listening & Speaking	Analytic scoring of individual item traits	Content Fluency Pronunciation Vocabulary Control Grammatical Accuracy
	Part 3	Multiple choice single answer	Listening	Dichotomously scored: correct/incorrect	
	Part 4	Present an aspect of a lecture	Listening & Speaking	Analytic scoring of individual item traits	Content Fluency Pronunciation
	Part 5	Debate: Take notes	Listening & Writing	Analytic scoring of individual item traits	Note-taking (lectures, seminars, meetings, etc.)
	Part 6	Debate: Express your opinion in writing	Listening & Writing	Analytic scoring of individual item traits	Content Grammatical accuracy Vocabulary control
WRITING	Part 1	Summarize written text	Reading & Writing	Analytic scoring of individual item traits	Content Grammatical accuracy Vocabulary control
	Part 2	Write essay	Writing	Analytic scoring of individual item traits	Content Grammatical accuracy Vocabulary control Vocabulary range Coherence and cohesion

6.2 Raters

All human raters who assess test taker responses have recognized English language-teaching qualifications as well as significant English language-teaching experience. Raters follow a standardized training and certification process before starting live marking. Their marking is then monitored to ensure accuracy and consistency over time.

Automated quality assurance monitoring is carried out using test taker responses which have been marked previously by a number of experienced raters and so have agreed

benchmark ratings. These benchmark responses are included together with the new test responses to check that the raters continue to mark accurately within set tolerances.

Raters whose inter- or intra- rater reliability 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 if completed successfully. Raters who do not successfully complete re-standardization are permanently withdrawn from marking.

6.3 Scoring rubrics

Analytic scoring is used for the assessment of speaking and writing skills, which includes the integrated item types of listening-into-speaking, listening-into-writing, reading-into-speaking and reading-into-writing. A separate score is awarded for each of the item traits (item-trait scoring), as opposed to one holistic score for the task.

This allows for more exact reporting of proficiency, especially where language skills may be developing at different rates. Item-trait scoring also leads to greater reliability as each test taker is awarded a number of scores. Raters are required to focus on each trait individually, thus ensuring that they are all addressing the same features of the test taker's performance.

Each rating scale for each item trait that is assessed ranks the performance along an ordinal scale of 0 to 6 and describes language proficiency as a series of levels against which the test taker's performance is judged.

Each level is characterized in terms of what the test taker can do with the language (functions which can be performed) and their control of linguistic features, e.g., vocabulary, pronunciation, fluency, etc. The development of the scoring rubrics is informed by the relevant CEFR descriptor scales.

6.3.1 Scoring Rubrics for Speaking

6.3.1.1 Rating Scales for Linguistic Features

Score	Fluency	Pronunciation	Vocabulary Control	Grammatical Accuracy
6	Test taker speaks with a natural, effortless, unhesitating flow.	Test taker has a high level of phonological control including over prosodic features such as word and sentence stress, rhythm and intonation - so that the finer points of the text are revealed clearly and precisely. Intelligibility is not affected in any way by features of accent that may be retained from other language(s).	Consistently correct and appropriate use of vocabulary.	Maintains consistent grammatical control of complex language.
5	Test taker speaks fluently, almost effortlessly.	Test taker employs the full range of phonological features in their response with sufficient control to ensure intelligibility throughout. Some L1 interference may be noticeable, but they do not affect intelligibility at all.	Occasional minor slips, but no significant vocabulary errors.	Consistently maintains a high degree of grammatical accuracy; errors are rare and difficult to spot.
4	Test taker speaks mostly fluently with few hesitations.	Test taker almost always uses appropriate intonation, places stress correctly and articulates individual sounds clearly; accent tends to be influenced by L1 but has no effect on intelligibility.	Lexical accuracy is high. Occasional errors do not hinder communication.	Good grammatical control; occasional "slips" or non-systematic errors and minor flaws in sentence structure may still occur, but they are rare and can often be corrected in retrospect.
3	Test taker speaks with a fairly even tempo. There may be some hesitations.	Test taker generally uses appropriate intonation, places stress correctly and articulates individual sounds clearly, accent tends to be influenced by L1, but has little or no effect on intelligibility.	Lexical accuracy is generally high, though some confusion and incorrect word choice do occur without hindering communication.	Has a good command of simple language structures and some complex grammatical forms. Does not make mistakes which lead to misunderstanding.
2	Test taker speaks with relative ease. There may be noticeable pauses.	Test taker is mostly intelligible; approximates intonation and stress at both utterance and word levels. Accent is likely to be influenced by L1.	Shows very good control of elementary vocabulary but errors still occur when expressing more complex thoughts or handling unfamiliar topics and situations.	Communicates with reasonable accuracy in familiar contexts; generally good control, though with noticeable L1 influence. Errors occur, but it is clear what they are trying to express.

Score	Fluency	Pronunciation	Vocabulary Control	Grammatical Accuracy
1	Test taker speaks with an uneven tempo and noticeable pauses and hesitations.	Test taker is generally intelligible; can approximate intonation and stress at both utterance and word levels. However, accent is usually influenced by L1. Mispronunciation of individual words or sounds reduces clarity at times.	Shows good control of elementary vocabulary but major errors still occur when expressing more complex thoughts or handling unfamiliar topics and situations.	Uses reasonably accurately a repertoire of frequently used "routines" and patterns associated with more predictable situations.
0	Test taker speaks with irregular phrasing and much pausing. There are very noticeable hesitation and false starts.	Pronunciation is generally clear enough to be understood, but mispronunciations are frequent and cause some difficulty for the listener. A strong influence from L1 on stress, rhythm and intonation may also affect intelligibility.	Can control a narrow repertoire of basic vocabulary which is not sufficient to deal with the task.	Uses some simple structures correctly but still systematically makes basic mistakes.

6.3.1.2 Rating Scales for Content (Functions)

Score	Answer interview questions	Describe image	Read aloud	Present an aspect of a lecture	Express your opinion
6	Test takers at this level can keep up their side as interviewee extremely well and effortlessly, answering the question comprehensively in a structured response.	Test takers at this level can interpret and describe clearly and reliably the empirical data or visually organized information in the image.	All words of the prompt are read out correctly.	Test takers at this level can follow the lecture with ease and produce a clear, smoothly flowing, well-structured response fully addressing the task.	Test takers at this level can develop a clear, smoothly flowing, well-structured argument with an effective logical structure that fulfills the task.
5	Test takers at this level can give clear, detailed responses expanding and developing their response fully with little obvious searching for expressions.	Test takers at this level can interpret and describe clearly and reliably most salient points of the empirical data or visually organized information shown in the image.	Almost all words of the prompt are read out correctly. The response contains up to 3 mistakes. Each omission, insertion or replacement of a word counts as an error.	Test takers at this level can follow the lecture with relative ease and can relay the relevant information clearly and in detail.	Test takers at this level can develop a well-structured argument highlighting significant points with supporting examples and concluding appropriately.
4	Test takers at this level can give an effective and sustained response without much sign of having to restrict what they want to say. May fail to address an aspect of the task.	Test takers at this level can interpret and describe clearly and reliably significant points and relevant supporting detail contained in the image. May fail to address an aspect of the image.	Most words of the prompt are read out correctly. The response contains up to 5 mistakes. Each omission, insertion or replacement of a word counts as an error.	Test takers at this level can follow most aspects and the point of view expressed in the lecture and give a clear, detailed account with appropriate highlighting of significant points. They may fail to address an aspect of the task.	Test takers at this level can develop an argument systematically with appropriate highlighting of significant points, and relevant supporting detail. They may fail to address an aspect of the task.
3	Test takers at this level can expand and develop ideas with some precision. May fail to address an aspect of the task.	Test takers at this level can interpret and describe reliably detailed information contained in the image. May fail to address an aspect of the image.	Most words of the prompt are read out correctly. The response contains up to 8 mistakes. Each omission, insertion or replacement of a word counts as an error.	Test takers at this level can follow most aspects of the lecture, distinguish main themes and recognize the point of view expressed. Can give a clear, detailed response. They may fail to address an aspect of the task.	Test takers at this level can develop a clear argument, expanding and supporting their points of view at some length. They may fail to address an aspect of the task.

Score	Answer interview questions	Describe image	Read aloud	Present an aspect of a lecture	Express your opinion
2	Test takers at this level can provide concrete information required to answer the interview questions with some confidence, but with limited precision. May fail to address several aspects of the task.	Test takers at this level can interpret and describe some aspects in detail even though lexical gaps may cause hesitation or imprecise formulation.	Most of the words of the prompt are read out correctly. The response contains up to 12 mistakes. Each omission, insertion or replacement of a word counts as an error.	Test takers at this level can follow and relay a specific aspect of the lecture distinguishing between main ideas and supporting details. They may fail to address several aspects of the task.	Test takers at this level can develop an argument well enough to be followed without difficulty most of the time and give simple reasons to justify a viewpoint. They may fail to address several aspects of the task.
1	Test takers at this level can make themselves understood and communicate their main ideas with a wide range of simple language. May fail to address several aspects of the task.	Test takers at this level can interpret and describe in simple sentences some aspects, even though lexical limitations cause difficulty with formulation at times.	Most of the words of the prompt are read out correctly. The response contains up to 15 mistakes. Each omission, insertion or replacement of a word counts as an error.	Test takers at this level can understand and reasonably fluently relay the main points of a lecture in a linear sequence of points. They may fail to address several aspects of the task.	Test takers at this level can express their opinions using simple expressions, briefly give reasons and explanations to justify their opinions. They may fail to address several aspects of the task.
0	Test takers at this level use a series of simple phrases and sentences to make themselves understood, but may fail to address most of the aspects of the task.	Test takers at this level can describe in simple sentences the main aspect shown in the image. Pauses, false starts and reformulation may be very evident.	The response contains more than 15 mistakes. Each omission, insertion or replacement of a word counts as an error.	Test takers at this level can outline basic information in simple terms, but may fail to address most of the aspects of the task.	Test takers at this level can present their opinion in simple terms, but might fail to address most of the aspects of the task.

6.3.2 Scoring Rubrics for Writing

6.3.2.1 Rating Scales for Linguistic Features

Score	Vocabulary Range	Vocabulary Control	Grammatical Accuracy	Coherence & Cohesion
6	Has a good command of a very broad lexical repertoire.	Consistently correct and appropriate use of vocabulary.	Maintains consistent grammatical control of complex language.	Creates coherent and cohesive text making full and appropriate use of a variety of organizational patterns and a wide range of cohesive devices.
5	Has a good command of a broad lexical repertoire allowing gaps to be readily overcome with circumlocutions.	Occasional minor slips, but no significant vocabulary errors.	Consistently maintains a high degree of grammatical accuracy; errors are rare and difficult to spot.	Produces clear, smoothly flowing, well-structured language, showing controlled use of organizational patterns, connectors and cohesive devices.
4	Has a good range of vocabulary; makes meaning clear in spite of some imprecisions.	Lexical accuracy is high. Occasional errors do not hinder communication.	Good grammatical control; occasional "slips" or non-systematic errors and minor flaws in sentence structure may still occur, but they are rare and can often be corrected in retrospect.	Produces clear and coherent text using a variety of linking expressions efficiently to mark clearly the relationships between ideas, although "jumpiness" may occur.
3	Has a good range of vocabulary; varies formulation to avoid frequent repetition, but lexical gaps are evident.	Lexical accuracy is generally high, though some confusion and incorrect word choice do occur without hindering communication.	Has a good command of simple language structures and some complex grammatical forms. Does not make mistakes which lead to misunderstanding.	Produces text that is generally well-organized and coherent, using a range of linking expressions and cohesive devices.
2	Has a good range of mainly basic vocabulary, which may not always be sufficient to deal with some aspects of the task.	Shows very good control of elementary vocabulary but errors still occur when expressing more complex thoughts or handling unfamiliar topics and situations.	Communicates with reasonable accuracy in familiar contexts; generally good control, though with noticeable L1 influence. Errors occur, but it is clear what they are trying to express.	Links a series of shorter, simple elements including counter arguments into a connected text which may be less well structured in places.
1	Has sufficient vocabulary to express themselves in simple terms, which may not always be sufficient to deal with some aspects of the task.	Shows good control of elementary vocabulary but major errors still occur when expressing more complex thoughts or handling unfamiliar topics and situations.	Uses reasonably accurately a repertoire of frequently used "routines" and patterns associated with more predictable situations.	Links a series of shorter, simple elements into a connected, linear sequence of points using a limited number of cohesive devices. The text may be less well structured and arguments may be poorly linked.

Score	Vocabulary Range	Vocabulary Control	Grammatical Accuracy	Coherence & Cohesion
0	Has insufficient range of vocabulary to deal with the task.	Can control a narrow repertoire of basic vocabulary which is not sufficient to deal with the task.	Uses some simple structures correctly, but still systematically makes basic mistakes.	Uses only a few simple connectors to link simple sentences, which are largely insufficient to produce a coherent line of argument.

6.3.2.2 Rating Scales for Content (Functions)

Score	Express your opinion	Summarize written text	Write essay
6	Test takers at this level can produce a clear, smoothly flowing response with an effective logical structure that fulfills the task.	Test takers at that level can summarize in writing all key aspects and supporting details producing a clear, well-structured and smoothly flowing summary.	Test takers at this level can produce a clear, smoothly flowing, complex essay with an effective logical structure which helps the reader identify significant points. They are able to convey finer shades of meaning by using, with reasonable accuracy, a wide range of qualifying devices.
5	Test takers at this level can produce a clear, well-structured response emphasizing the salient points; expand and support points of view with subsidiary points, reasons and relevant examples.	Test takers at that level can summarize in writing the salient points, interpreting the content appropriately and producing a coherent summary.	Test takers at this level can produce a clear, well-structured essay emphasizing the relevant salient points and expand and support points of view at some length with subsidiary points, reasons and relevant examples.
4	Test takers at this level can produce a text which develops an argument systematically with appropriate highlighting of significant points and relevant supporting detail. They may fail to address an aspect of the task.	Test takers at that level can summarize in writing the main content and include supporting inference made with reference to specific information in the text.	Test takers at this level can produce an essay which develops an argument systematically with appropriate highlighting of significant points and relevant supporting detail. They may fail to address an aspect of the task.
3	Test takers at this level can produce a text which develops an argument and provides reasons supporting a particular point of view, though their language lacks expressive power. They may fail to address an aspect of the task.	Test takers at that level can summarize in writing the main content.	Test takers at this level can produce an essay which develops an argument and provides reasons supporting a particular point of view, though their language lacks expressive power. They may fail to address an aspect of the task.
2	Test takers at this level can produce a short response using simple language to give and justify their opinion with reasonable precision. They may fail to address several aspects of the task.	Test takers at that level can summarize in writing some of the main points contained in the text.	Test takers at this level can produce a short text using simple language to give and justify their opinion with reasonable precision. They may fail to address some aspect of the task.
1	Test takers at this level can produce a simple response, which states their main opinion and some reasons. They may fail to address several aspects of the task.	Test takers at that level can summarize in writing some of the main points using the original text wording and ordering.	Test takers at this level can produce a very brief text getting across the main point they want to make comprehensively. They may fail to address an aspect of the task.
0	Test takers at this level can communicate what they want to say in a simple response with limited information; generally, the message is	Test takers at that level can pick out short phrases from which are insufficient to provide an appropriate summary.	Test takers at this level can produce simple texts, linking sentences with connectors like "and", "because" or "then".

compromised and the test takers fail to address most of the aspects of the task.		OR: The essay contains fewer than 140 words.

7 Alignment to the CEFR

The CEFR is recognized globally as a key framework for interpreting language proficiency. Many institutions base materials, teaching programs and tests on the CEFR levels. The CEFR includes a set of language levels defined by descriptors of language competencies. In developing GATEWAY, its alignment to the CEFR was paramount. Consequently, the CEFR has been embedded in the test development process as follows:

7.1 Employing CEFR Can Do statements in the test design

Careful attention was given to the alignment of the test specifications, item–writing guidelines and the test items to targeted CEFR Can Do statements that either focus on linguistic features or language functions. The test content is specifically designed to elicit performances at the CEFR proficiency levels A2–C2. Test takers who score below level A2 receive the result *below A2*. This indicates that they are below the levels reported in the test and no specific CEFR level can be assigned to their performance.

Item writers, reviewers and raters underwent extensive training in interpreting and applying the CEFR correctly to ensure the writing and rating of test items remains rooted in the framework.

7.2 Analyzing performance of field-tested items

The test has been field tested globally with over 1,000 students across 50 countries from a wide number of first-language backgrounds at each of the targeted CEFR levels. Data related to overall item quality, test taker performance and the extent to which test items could be scaled to the intended CEFR levels was collected. Using Rasch analysis the difficulty of each item was plotted based on anchor items. Inferences could then be made about the CEFR levels of each item.

8 Test delivery

The GATEWAY test is an online test that can be delivered anywhere at any time. The item bank contains a large number of items of each item type. The test operates by using this item bank and a series of item selection rules to assign a unique version of the test to each test taker. This means GATEWAY does not have fixed test versions in which all test takers encounter the same set of questions. This significantly increases test security.

8.1 Registration

Test takers register online through english3.com. The test can be taken on demand.

8.2 Test Taking Requirements

GATEWAY is delivered online via the English3 learning management system. To complete the test, test takers need:

1. A laptop or desktop computer with a microphone, speakers, and a front-facing camera.
2. A passport, driver's license, or other government-issued photo identification. The name on their ID must fully match the name they register with. Test takers will have the opportunity to review and correct any information prior to beginning the test.
3. A well-lit and quiet space where they can complete the test entirely on their own.
4. A reliable internet connection. Their internet speed is tested during the test setup phase and test takers are prompted to use a faster internet connection if the one they are using does not meet minimum requirements.
5. Chrome browser to implement security features.
6. Up to 2 hours to complete the test in one sitting. A timer is included on all screens once the test taker begins the test.

Test takers need to obey the following rules:

1. Do not click away from the page containing your GATEWAY test. Any attempt to do so is considered malpractice and may revoke your test session.
2. Do not use any copy, cut, or paste function with your mouse, keyboard, voice command, or device. Any attempt to do so is considered malpractice and may revoke your test session.
3. Do not take screenshots, record the screen or otherwise attempt to take any of the test content. Any attempt to do so is considered malpractice and may revoke your test session.
4. Give consent to sharing your screen as well as to monitoring your test session via webcam and microphone. This is mandatory throughout the whole test to ensure the integrity of the GATEWAY qualification and results.

Test takers are made aware that any malpractice will be investigated: If any malpractice is suspected, the test will be flagged as suspicious. In order to maintain the integrity of the GATEWAY qualification and results, all allegations of malpractice will be investigated. Depending on the severity of the violation, English3 might be required to revoke the suspicious test session.

Before starting the test, test takers are required to sign an Academic Honesty Agreement declaring that they will adhere to the test taking rules.

8.3 Test Security

The GATEWAY test has profited from our expertise gained when delivering other online testing products including the English3 J-1 Interview and E3PT. The Technical Advisory Group that consists of language testing experts as well as our stakeholders in the admissions process have informed and supported the development of robust test security features.

English3 enacts both technology- and human-driven security methods to ensure the integrity of the GATEWAY test. Online testing allows for a far more dynamic and holistic view of test taker ability. Innovative item types are assessing test taker proficiency in more effective and efficient ways as we put the test taker into an environment that is very similar to the academic setting in which they will be expected to communicate. GATEWAY simulates, for example, lectures and group discussions.

At English3 we are layering in security measures to address the primary challenges in the testing industry of validity, reliability, and security. In general, these layers include computer-based security flagging systems coupled with human evaluation at differing levels. In particular, we utilize asynchronous video monitoring, confirmation of ID, facial recognition, plagiarism recognition software, automated prevention of malpractice, etc.

As a test provider, English3 must be confident that the score report is an accurate representation of the test taker's proficiency in English. The test security focus is on preventing test takers from gaining an unfair advantage. Our security measures focus on deterrents or preventative measures for each threat. Content theft, pre-knowledge, proxy test takers, and malpractice while taking the test are among the top threats to test security.

Test takers are asked to sign an Academic Integrity Agreement in which they declare to follow the rules of honest test taking. For GATEWAY, we utilize facial recognition software to compare the likeness of the ID photo and the test taker. Test takers are being recorded throughout the test.

They may take on-screen notes in areas indicated during the test. These notes can be reviewed by proctors. Any unusual behavior, for example, looking away from the screen for longer than a natural glance; a second person appearing, or attempts at copying the screen is automatically flagged and in certain cases the test taker is prevented from continuing with the test.

Once a test is submitted data forensics are used to detect any issues such as large score discrepancy between extended and closed responses which could indicate that items have potentially been exposed and need replacing.

As GATEWAY has been developed as an online test, we also ensured that the item types are fit for purpose. Most tasks require a real-time extended response. English3's proprietary technology assesses real-time ability to produce language reducing security concerns including gaming and content theft.

The GATEWAY test is graded partly by computer (dichotomously scored responses) and partly by trained raters (extended responses). As tests are looked at by human raters, security increases as any concerns will be investigated. Multiple-choice or gap-fill tasks have been minimized not only to increase the security of GATEWAY but also to achieve a more accurate assessment of test takers' abilities to utilize the language at university.

To further increase security, English3 works with remote proctors asynchronously. Proctors are long-term contractors who have been vetted before employment. Their contract explains their ethical obligations and consequences if found in breach of any contractual obligations. Tests are reviewed by proctors, who are able to see the responses as well as review the video of the test taker while responding to each question.

As mentioned above, when malpractice is suspected, test takers might have their test revoked, they would also lose the funds they used to register for the test, and they have a mandatory waiting period imposed. To avoid such negative consequences, great effort is taken in establishing and maintaining proper test security. This is reflected in the course of actions outlined below.

1. Mitigating the risks of cheating
    a. By providing a secure testing environment
        i. Outlining clear expectations of the test taker by having test takers review and agree to testing policies prior to beginning their test.
        ii. Displaying their webcam video on all test pages.
        iii. Including a timer on all pages.
        iv. Implementing a restricted user interface with prompts keeping the test taker on the test window.
    b. Through test design and delivery
        i. Integrated item types
        ii. Item banking
            1. Test items are selected by stratified random sampling at the beginning of the test.
        iii. Item types requiring spontaneous open responses with limited preparation time given.
2. Addressing common forms of cheating:
    a. Test administrator / test taker collusion
        i. Not possible with the online delivery format of GATEWAY.
    b. Counterfeit score reports
        i. Every institution will access Score Reports on the English3 website through their institution's online portal. We recommend all student scores are reviewed on the university's English3 dashboard
    c. Proxy test takers
        i. Test takers upload an approved government-issued photo ID. Their photo ID will be compared with their test video.
    d. In-room cheating
        i. GATEWAY maintains a one-to-one test taker to proctor ratio. The entirety of the test video is reviewed by trained proctors and AI-based software.
    e. Content harvesting, leaks, or theft

    i. Our Chrome extension prevents screenshots, copy/paste, and many other tactics used to save material. Web crawlers scan the internet for leaked content.
 3. Proctoring supported by technology
   a. Data forensics technology identifies unusual test-taker behavior. Suspicious activities are flagged and then reviewed by proctors.
   b. Checking test-taker equipment:
     i. Webcam video
     ii. Audio
     iii. The screen of the test taker
   c. Chrome browser functionality
     i. Offers continuous screen monitoring.
     ii. Disables copy/paste.
     iii. Ensures no other screen recording programs are running.
     iv. Disables ability to continue if a second monitor is attached
     v. Disables ability to open new tabs.
     vi. Disables ability to leave the test window.

8.4 Results reporting

Test takers receive their scores within 5 days, or within 24 hours, depending on the test option purchased. Test takers are able to submit test scores to an unlimited number of institutions through their English3 account.

Scores are considered valid for 2 years following the date the test was completed. The nature of acquiring a new language and the natural process of language atrophy require the test score validity to expire.

Institution administrators receive access to an online portal where they can listen to the test taker's video introduction, see their photo ID and demographics and verify test taker scores.

The image below shows an example of the GATEWAY score report test takers will be able to send or download. The report shows the Overall score as well as the Communicative Skills scores, i.e., speaking, listening, reading and writing together with the respective CEFR equivalents. A description of the level of English language proficiency is also provided.

8.5 Results reviews and appeals

Test takers may request a rescore of their most recent GATEWAY attempt by contacting support@english3.com. The rescore fee is $49. If the overall score or a subscore is incorrect, the rescoring fee will be refunded. Only spoken and written responses are rescored. If the score changes (up or down), it will replace the original score.

GATEWAY

English3
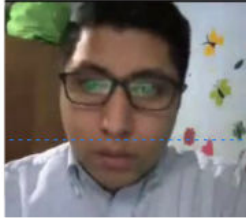
Name	Moises Vazqez
Email	mvasqez89@gmail.com
Applicant ID	4568845484654845
Test Date	Mar 12, 2022
Score Issued Date	Mar 15, 2022
Country of Residence	Mexico
First Language	Spanish

Overall score

510 C1 CEFR

Skill scores

495 Speaking   **525** Listening   **520** Reading   **500** Writing

Test takers at this level understand standard language on both familiar and unfamiliar topics encountered in personal, social, academic or vocational life. They can follow the essentials of lectures, talks and reports and other forms of academic/professional presentation which are propositionally and linguistically complex. They are able to give clear, systematically developed descriptions and presentations, with appropriate highlighting of significant points, and relevant supporting detail. They can produce an essay which develops an argument with appropriate highlighting of significant points and relevant supporting detail.

Learn more at english3.com/scores

The GATEWAY test is a comprehensive academic English assessment. Institutions should validate all scores through their online dashboard.

Applicants can log into elearn.english3.com to send official Score Reports to additional institutions. Scores are valid for two years from the Test Date. Applicants may take the GATEWAY test up to two times per months. Please email info@english3.com with questions.

English3

9 Accessibility

See up-to-date accessibility information at https://english3.com/english3-proficiency-test/english3-disability-accomodations/.

English3 is committed to providing reasonable testing accommodations for test takers with documented disabilities or health-related needs, as recognized under the Americans with Disabilities Act Amendments Act (ADAAA). Requests for accommodations are reviewed and responded to within two weeks.

When to Register: Once a test taker is approved to receive accommodations, English3 will provide a unique registration link. Wait to register until you receive accommodation approval.

Accommodation Options

- Extended testing time. Time and a half (1.5x) and double time (2x)
- Listening section omitted
- Sign language interpreter
- Writer/recorder of answers
- Speaking section omitted
- Additional breaks
- Transcripts of audio elements in Speaking and Writing sections

Required Documents

- State diagnosed disability or disabilities and describe the functional limitations resulting from the disability or disabilities.
- Include educational, developmental and medical history relevant to the disability for which testing accommodations are being requested.
- List all test instruments used in the evaluation report to document the stated disability.
- Submit on official letterhead and be signed by an evaluator qualified to make the diagnosis, including license or certification in the relevant area of specialization.
- Be current. Current means:
- last six years for attention deficit/hyperactivity disorder (ADHD), learning disabilities (LD), autism spectrum disorder (ASD), or intellectual disability (ID)
- last two years for psychiatric disabilities and traumatic brain injury
- *Note: This requirement does not apply for permanent or unchanging disabilities.
- Describe the specific accommodations requested.
- Adequately support each of the requested testing accommodation(s).
- Note: All required documents must be submitted in English or in certified translation

If test takers plan to request accommodations other than those listed above, they may note those accommodations in the Additional Information section of the request form.

Confidentiality: English3 maintains the confidentiality of your documentation and will not release any part of the information you provide without your informed consent.

10 References

Biber, Douglas (et al.). *Representing Language Use in the University: Analysis of the TOEFL 2000 Spoken and Written Academic Language Corpus.* Princeton, New Jersey: ETS®, 2004.

Biber, Douglas. *University Language: A corpus-based study of spoken and written registers.* Amsterdam: John Benjamins B.V., 2006.

Chalhoub-Deville, Micheline & Barry O'Sullivan. *Validity: Theoretical Development and Integrated Arguments.* Volume 3 of British Council Monographs on Modern Language Testing. Equinox Publishing Limited, 2020.

Chamot, Anna Uhl & Michael O'Malley. *The CALLA handbook: Implementing cognitive academic language learning approach.* Reading, MA: Addison-Wesley, 1994.

Collier, Virginia P. "Age and rate of acquisition of second language for academic purposes." *TESOL Quarterly* 21 (1987): 617–641.

Council of Europe (et al.). Commo*n European Framework of Reference for Languages: Learning, teaching, assessment.* Cambridge: CUP, 2001.

Council of Europe. *Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR) - A Manual.* Strasbourg: Council of Europe Publishing, 2009.

Council of Europe. Commo*n European Framework of Reference for Languages: Learning, teaching, assessment - Companion volume.* Strasbourg: Council of Europe Publishing, 2020.

Davies, A. (et al.). *Dictionary of language testing.* Cambridge: CUP, 1999.

Douglas, Dan. *Assessing Languages for Specific Purposes.* Cambridge: CUP, 2000.

Luoma, Sari. *Assessing Speaking.* Cambridge: CUP. 2004.

McKay, Sandra. *Teaching English as an International Language: Rethinking Goals and Approaches.* Oxford: OUP, 2002.

McNamara, Tim. *Language Testing.* Oxford: OUP 2000.

Morrow, Keith (ed.). *Insights from the Common European Framework.* Oxford: OUP, 2004.

Nation, I.S. Paul. *Teaching and Learning Vocabulary.* New York: Newbury House, 1990.

Pearson Longman. *Official Guide to Pearson Test of English Academic.* Hong Kong: Pearson Longman Asia ELT, 2010.

Read, John. *Assessing Vocabulary*. Cambridge: CUP, 2000.

Scarcella, Richard C. et al. (eds.). *Developing communicative competence in a second language*. New York: Harper & Row, 1990.

Xi, Xiaoming & John M. Norris (eds.). *Assessing Academic English for Higher Education Admissions.* New York: Routledge, 2021.

Contributors

Kirsten Sutton, Test Development Manager, English3
Anthony Celentano, Assessment Specialist
Helen Wilson, English Language Specialist
Moroni Flake, CEO, English3