

CAEL CE 2018 Annual Report.

*Paragon Testing Enterprises
110-2925 Virtual Way, Vancouver, BC V5M 4X5
research@paragontesting.ca*



CAEL[®]

Table of Contents

1 Description of the Test	4
1.1 General Description	4
1.2 Test Format	4
2 Scoring and Reporting of Results	5
3 Interpreting CAEL Results	5
4 Test-Taking Population	6
5 Test Statistics	7
5.1 Descriptive Statistics and Distribution of Component Scores	7
5.2 Measurement Consistency	8
5.3 Rater Agreement for the Writing and Speaking Components	8
Appendices	10
Appendix A CAEL CE Reporting Scale	10

List of Tables

Table 1: Format and Content of the CAEL CE Test	4
Table 2: Interpretation of CAEL CE Test Scores	5
Table 3: Distribution (in %) of CAEL CE Test Takers by Purpose for Taking the Test	7
Table 4: Descriptive Statistics for the CAEL CE Test	7
Table 5: Distribution (in %) of Test Takers at each Band Score	7
Table 6: Internal Reliability Estimates for the Multiple-Choice Forms on the CAEL CE Test	8
Table 7: Rater Agreement	9

1 Description of the Test

1.1 General Description

The Canadian Academic English Language Computer Edition (CAEL CE) Test is a standardized test designed to measure the English language proficiency of students planning to study in Canadian post-secondary institutions. It is administered by Paragon Testing Enterprises Inc., a subsidiary of The University of British Columbia.

The CAEL CE Test provides an authentic representation of language use in a Canadian academic context. As would be expected in a first-year Canadian university or college classroom, test takers read articles, listen to a lecture, answer questions, and write a short essay. Each test taker receives a score report showing their performance on each component as well as an overall score that is the unweighted average of the four individual component scores. CAEL CE scores are accepted by more than 180 academic institutions across Canada as proof of the English language proficiency of applicants. A number of professional associations also recognize CAEL CE as proof of English language proficiency required for membership.

Paragon is committed to upholding the highest standards in educational measurement. All parts of the CAEL CE Test are written following specified guidelines, and results are closely monitored to ensure that they are accurate and informative. Paragon works closely with test centres to make certain that the CAEL CE Test is administered in a way that is fair and accessible but also secure, ensuring that the CAEL CE Test is available to all individuals who wish to take the exam and that the scores issued are defensible.

1.2 Test Format

CAEL CE assesses test takers' English language proficiency in an academic context. The test reports scores on four components—Speaking, Reading, Listening, and Writing. Test takers complete a range of tasks. Some of these tasks will require test takers to use what they have read and listened to in order to answer a question in speaking or writing. Table 1 describes the format and content of each test component.

Table 1: Format and Content of the CAEL CE

Component	Time Allotted	Component Sections	Number of Tasks
Speaking	7-10 minutes	Two speaking tasks, each based on a short question One speaking task based on a graph/diagram/chart	3
Integrated Reading	35-50 minutes	One or two short reading passages with comprehension questions One or two long reading passages with comprehension questions One speaking question, answered using material from a long reading passage	15 - 26
Integrated Listening	25-35 minutes	One or two short listening passages with comprehension questions One or two long listening passages with comprehension questions One speaking question, answered using material from a long listening passage	15 - 26
Academic Unit A	60-70 minutes	One long reading passage with comprehension questions One long listening passage on the same topic, with comprehension questions One writing question requiring an extended response, using material from both the long reading passage and the long listening passage	23 - 31
Academic Unit B	40-45 minutes	One long reading passage with comprehension questions One long listening passage on the same topic, with comprehension questions One writing question requiring a short response, using material from the long reading passage or the long listening passage	23 - 31

*Unscored Items: Each test contains unscored items used for test development. These unscored items can be found anywhere within each test and have the same format as the scored items.

2 Scoring and Reporting of Results

The multiple-choice items are scored by computer. Each correct answer contributes proportionately to the final score, and there are no points deducted for wrong answers. The Speaking and Writing components are evaluated by at least two Paragon-certified raters according to a scale established by Paragon.

Test takers receive a score report including their individual English proficiency scores for the Speaking, Reading, Listening, and Writing, components, as well as an overall score. CAEL scores are reported on a scale from 10 – 90 with accompanying descriptors of what performance represents. The overall score is calculated as an equally weighted average of the four component scores rounded to the nearest ten-point interval. Table 2 presents descriptions of test taker proficiency at each full band level.

Table 2: Interpretation of CAEL CE Test Scores

CAEL CE Band	Descriptor
80-90	Expert: Demonstrates a high level of competence, accuracy, and effectiveness in academic/professional settings
70	Adept: Uses generally accurate language in most settings; some limitations in flexibility are evident
60	Advanced: Displays competence in academic or professional settings
50	High Intermediate: Exhibits some competence in academic or professional settings; communication may break down in places
40	Intermediate: Demonstrates some ability to comprehend and articulate complex ideas and arguments typical of academic or professional settings
30	High Beginner: Expresses basic ideas about familiar topics in routine settings
10-20	Low Beginner: Communicates with limited ability

For a more detailed review of CAEL score interpretation, please see Appendix A.

3 Interpreting CAEL CE Results

The CAEL CE has been designed to assess the English language ability of test takers in academic contexts. Table 2 (previous section) describes the interpretation of each band score.

When interpreting a CAEL CE Test score report, it is important to remember that the CAEL CE estimates test takers' true proficiency by approximating the kinds of tasks that they may encounter in the course of their study or academic work. There are, inevitably, small mismatches between the tasks that test takers complete as part of the test and the tasks that they have to perform in a specific context. Also, temporary factors unrelated to test takers' true proficiencies, such as fatigue, anxiety, or illness, may affect their CAEL CE results.

When using test scores for decision-making, check the date the test was taken. Language ability changes over time; it can improve with active use and further study of the language, but may also diminish if an individual does not continue to study or use English on a regular basis. It is also important to remember that test performance is only one aspect to be considered. Communicative language ability consists of both knowledge of language and knowledge of the world. Therefore, it is important to consider how factors other than language affect how well someone can communicate. For example, in an academic context, the ability to function effectively involves not only knowledge of English but also other knowledge and skills such as prior knowledge, subject matter expertise, and the ability to integrate knowledge from multiple sources.

4 Test-Taking Population

This section presents an overview of the test takers who took the CAEL CE Test in 2018, providing demographic information for the test population.

Table 3 shows that the majority of CAEL CE test takers registered with the intent to use their test score as proof of English language proficiency for post-secondary applications.

Table 3: Distribution (in %) of CAEL CE Test Takers by Purpose for Taking the Test

Purpose	% of Test Takers
For post-secondary admissions	54.6%
Employment	35.6%
Other	9.8%

5 Test Statistics

5.1 Descriptive Statistics and Distribution of Component Scores

Tables 4 and 5 present descriptive statistics for each component of the CAEL CE. The tables indicate that, overall, test takers score more highly on average on the Reading and Listening components than on the Speaking and Writing components, and that there is also more variation among scores for Reading and Listening than for Speaking and Writing.

For most universities which accept the CAEL CE as proof of English language proficiency, the minimum required score is an Overall 60, often with the additional requirement that no individual component score is less than 50. Although the mean Overall score is 59.95%, Table 5 shows that 65.6% of test takers achieve an Overall score of 60 or above, thus meeting the requirements for most university admissions.

Table 4: Descriptive Statistics for the CAEL CE Test

Component	Minimum Score	Maximum Score	Mean	Standard Deviation
Speaking	10	90	59.45	13.35
Reading	10	90	59.22	19.25
Listening	10	90	62.01	17.19
Writing	10	90	53.46	11.84
Overall	10	90	59.95	13.600

Table 5: Distribution (in %) of Test Takers at each Band Score

Band	Speaking	Reading	Listening	Writing	Overall
10	0.2%	1.8%	1.2%	1.4%	0.2%
20	0.6%	1.9%	1.8%	1.0%	1.0%
30	4.6%	3.5%	2.3%	2.6%	2.4%
40	7.0%	19.0%	9.2%	14.2%	9.0%
50	21.9%	17.5%	20.2%	32.2%	21.8%
60	28.3%	18.1%	19.4%	37.6%	29.1%
70	29.0%	13.2%	22.1%	8.5%	22.8%
80	6.5%	13.6%	14.5%	2.2%	12.3%
90	1.9%	11.4%	9.2%	0.3%	1.4%

5.2 Measurement Consistency

Test scores are a numerical measure of a test taker’s ability. In theory, a test taker’s score should be the same each time the test is taken. In practice, however, even when the test conditions are carefully controlled, an individual’s performance on a set of test items will vary from one administration to another due to variation in the items across different versions of the same test (This source of variation is mitigated through IRT form equating) or due to variability in individual performance. Among the reasons for variations in individual performance are temporary factors unrelated to a test taker’s English language proficiency. These include fatigue, anxiety, or illness. Therefore, test scores always contain a small amount of measurement error. The aim is to keep this error to a minimum. For standardized examinations such as the CAEL Test, a reliability figure of 0.80 and above is expected and acceptable.

Paragon monitors two facets of measurement consistency: internal consistency and test-retest reliability¹. Internal consistency is a measure of whether items in the test that are intended to measure the same construct produce similar results. Internal consistency estimates (Cronbach’s alpha) are calculated for each CAEL CE Listening and Reading form. In 2018 the mean reliability estimate for the Reading forms was 0.84 and the mean reliability estimate for the Listening forms was 0.82 (Table 6). These values suggest good internal consistency of measurement for the CAEL CE Reading and Listening component test forms.

Table 6: Internal Reliability Estimates for the Multiple-Choice Forms on the CAEL CE Test

Component	Mean Reliability	Standard Deviation
Reading	0.84	0.04
Listening	0.82	0.05

Table 6 also shows that the means among the Reading and Listening forms administered in 2018 were within the acceptable range, and the standard deviation figures indicate that few, if any individual forms have dipped below the 0.80 threshold. These values suggest good consistency of measurement both within and between each CAEL CE Reading and Listening form.

5.3 Rater Agreement for the Writing and Speaking Components

The raters for the Speaking and Writing components of the CAEL CE Test are highly proficient in English and are fully trained and certified by Paragon. Each of the test takers’ speaking and writing tasks are rated independently by two accredited raters. If the scores awarded by the raters fall outside an acceptable tolerance, the task is evaluated by a third rater.

¹ Test-retest reliability is the subject of ongoing research by Research & Development at Paragon Testing Enterprises. Test-specific research reports are made available at: <https://www.paragontesting.ca/about-research/test-reports/cael-test-reports/>

Paragon monitors rater agreement for quality control purposes. The percentage of within tolerance agreement between the raters is monitored. Since the evaluation process for the Speaking and Writing components relies on human judgements and the interpretation and application of a rating scale, variations in judgements are to be expected. Table 7 shows the rater agreement for the Speaking and Writing components in 2018 and shows high consistency of judgement between raters. The lower level of agreement for ratings on the Writing is matched by an increased proportion of third ratings, in order to ensure fairness and accuracy in the creating of all test taker scores.

Table 7: Rater Agreement

Component	Proportion in agreement
Speaking	0.87
Writing	0.77

Appendices

Appendix A CAEL CE Reporting Scale

Band Score	Reading Performance	Listening Performance	Writing Performance	Speaking Performance
80-90 Expert - Fluent	Reads academic texts with ease demonstrating comprehension equal to that of experienced academic readers.	Comprehends lectures as well as an experienced academic listener	Writes with authority and style demonstrating mastery of appropriate, concise and persuasive academic writing.	Speaks with authority on a variety of topics. Demonstrates flexibility, controls nuance and speaks with spontaneity and comprehensibility.
70 Adept	Reads and comprehends academic texts with ease. Can interpret information with flexibility.	Understands lectures with apparent ease.	Readily responds to the demands of the topic and presents information clearly and logically.	Speaks with ease presenting information clearly and logically.
60 Advanced	Understands main ideas and is able to identify most relevant details. Can interpret information with some flexibility.	Understand information regarding both main ideas and supporting details. May lack some flexibility and miss some information.	Can develop a thesis using a range of support and uses language that is generally accurate.	Speaks fluently, flexibly and with a degree of ease. Compensates strategically for limitations but communicates most required information clearly.
50 High Intermediate	Understands main ideas but may misinterpret information. Can identify some relevant details but reads slowly and with greater effort than most academic readers.	Processes most of the lecture for general ideas but may miss or misinterpret details. Overall understanding is still somewhat restricted.	Addresses the topic to a degree but with limited clarity and cohesiveness.	Speaks with some fluency and flexibility but speaks unevenly; at times there is a natural easy quality to the response but at other times the response breaks down.
40 Intermediate	Understands main ideas but is restricted by limited vocabulary and a lack of familiarity with textual conventions.	Identifies the meaning of some unfamiliar terms but overall understanding is restricted and uneven.	Makes links among ideas and addresses the topic but writing lacks clarity and cohesiveness.	Can speak with some fluency but without flexibility and with noticeable effort.
30 High Beginner	Understands some of the main ideas but is unable to identify specific, relevant ideas; reads with limited accuracy and fluency.	Makes some sense of sections of lectures by guessing but overall understanding is limited.	Writes something related to the topic but writing is not predictable and language is restricted.	Can speak but with false starts, hesitations and some mispronounced words. Provides studied and careful responses.
10-20 Low Beginner	Understands the main idea at times but misses almost all supporting detail.	Takes some meaning from individual words but overall understanding is sketchy and random.	Uses words randomly but language is very restricted and/or ungrammatical.	Communicates some information but mispronounces many words and speaks with great difficulty and many long pauses.