CAEL Annual Report of 2020 Test Takers

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

CAEL®

1 Description of the Test

1.1 General Description

The Canadian Academic English Language (CAEL) 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.

The CAEL 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 using information from the listening and reading materials. CAEL scores are accepted by more than 180 academic institutions across Canada as proof of an applicant's English language proficiency. A number of professional associations also recognize CAEL as proof of English language proficiency required for membership.

CAEL is fully computer-delivered and available via two methods of administration – CAEL at a Test Centre, and CAEL Online. CAEL Online includes the same test format, content, and reporting scale as the CAEL Test that is delivered at Paragon's test centres. The CAEL Test is available at test centres over 40 locations worldwide, with test dates available every month. CAEL Online can be safely taken from home and is overseen by a secure, online proctoring solution.

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

1.2 Test Format

CAEL assesses test-takers' English language proficiency in an academic context. Test takers complete a range of tasks in five test units. To incorporate the integrated nature of academic language skills, 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 unit.

Table 1: Format and Content of the CAEL Test

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.

1.3 Scoring and Reporting of Results

Each CAEL test taker receives a score report showing their performance on the Listening, Reading, Writing, and Speaking components, and an Overall score that is the unweighted average of the four component scores.

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 four Paragon-certified raters according to a scale established by Paragon.

CAEL scores are reported on a 9-band scale from 10 to 90 with accompanying descriptors of what the performance represents. The overall score is calculated as an average of the four component scores rounded to the nearest band level. Table 2 presents descriptions of test taker proficiency at each band level.

Table 2: Proposed Interpretation of CAEL Overall Band Scores

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

For a more detailed review of CAEL score interpretation, please see Appendix A: CAEL CE Reporting Scale.

1.4 Interpreting CAEL Results

CAEL 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 score report, it is important to remember that CAEL 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 test takers completed as part of the test and the tasks that they need 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 results.

It is also important to remember that test performance reflects only one aspect that impacts how well the candidate will perform in an academic context. Communicative language ability consists of both knowledge of language and knowledge of the world. In an academic context, the ability to function effectively involves not only knowledge of English but also knowledge of the context, subject matter expertise, and the ability to integrate information from multiple sources.

2 Test-Taking Population

This section presents an overview of the CAEL test taker population in 2020 in terms of test purpose, mode of administration, and demographic information.

Figure 1 shows the distribution of test takers by test purpose. The majority of CAEL test takers registered with the intent to use their test scores as proof of English language proficiency for either employment or post-secondary applications.

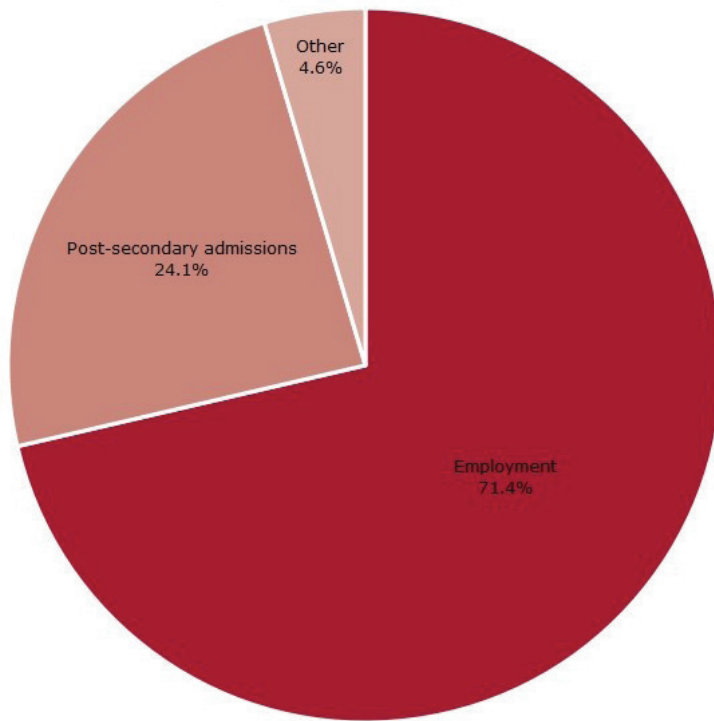**Figure 1: Distribution of CAEL Test Takers by Purpose**

Figure 2 shows the proportions of test takers by the two modes of administration. A slightly larger number of test takers took CAEL at a test centre.

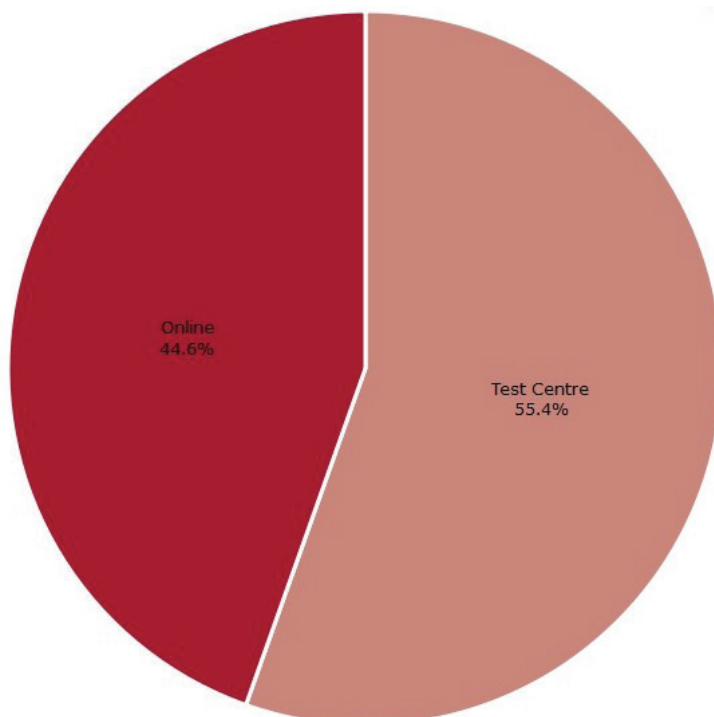**Figure 2: Online vs. Test Centre Population**

Table 3 shows that the test-taker population represents a wide range of age groups. Most of the test takers are in the age range of 26 to 45, representing 69.87% of all test takers in 2020.

Table 3: Distribution of CAEL CE Test Takers by Age (%)

Age Group	% of Total Test Population
< 20	4.99
21-25	9.00
26-30	17.13
31-35	20.70
36-40	18.86
41-45	13.18
46-50	8.55
51-55	4.60
> 55	3.01

3 Test Statistics

3.1 Score Distributions

This section presents a summary of the score distributions of each CAEL component and overall band levels. Table 4 presents the mean score and standard deviation for each component of CAEL. Table 5 shows the percentage distribution of test takers at each band level.

Table 4: CAEL Component and Overall Test Scores

Component	Mean	Standard Deviation
Listening	67.03	15.00
Reading	62.76	18.02
Speaking	61.11	9.82
Writing	66.34	9.61
Overall	65.59	11.31

Table 5: Distribution of CAEL Test Takers by Band Score (%)

Band	Listening	Reading	Speaking	Writing	Overall
10	0.50	0.98	0.17	0.33	0.00
20	0.53	1.20	0.19	0.19	0.53
30	1.20	2.26	0.89	0.64	0.75
40	4.21	13.96	1.48	2.23	2.34
50	14.60	15.21	6.77	14.79	11.00
60	20.56	20.84	25.21	56.82	31.23
70	27.08	17.53	54.57	16.49	34.40
80	19.61	14.49	9.30	7.27	17.72
90	11.73	13.54	1.42	1.23	2.03

For most universities that accept CAEL as proof of English language proficiency, the minimum required score is an Overall 60. As Table 5 indicates, 85.38% of test takers achieve an Overall score of 60 or above, thus meeting the requirements for most university admissions.

3.2 Measurement Consistency

In statistics and psychometrics, measurement consistency is often referred to as reliability. It is an important component when evaluating the quality of a test. A reliable test gives us the same result consistently, assuming no change in the individual's ability. For example, if a test is designed to measure English language proficiency and has been administered to the same individual multiple times, the test scores should be approximately the same if the test taker has not significantly improved their English proficiency during the period of the time. In contrast, an unreliable test produces inconsistent results each time, which greatly limits the value of the test scores.

Notably, even when the test conditions are carefully controlled, a small amount of measurement error is expected. An individual's performance on a set of test items will vary from one administration to another due to variability in individual circumstances, or due to variation in the items across different versions of the same test (this source of variation is mitigated through IRT form equating). High reliability of the test scores also indicates that the measurement error is kept at a minimum level.

Reliability can be estimated in multiple ways. In general, a higher value suggests greater reliability of the test scores. According to the literature and industry standards, a reliability of 0.80 and above is expected and acceptable for standardized examinations such as the CAEL Test.

3.2.1 Internal Consistency for the Listening and Reading Components

Internal consistency is a measure of whether test items designed to measure the same construct produce similar results. It is suitable for quantifying the reliability of tests that consist of many items, such as the Listening and Reading components on the CAEL Test. To measure internal consistency, we calculated Cronbach's alpha for each CAEL Listening and Reading form that has been administered to a minimum of 50 test takers in 2020.

Table 6: Internal Consistency for the CAEL Listening and Reading Forms

Component	Average Reliability	Mean Distance from the Median
Listening	0.80	0.03
Reading	0.82	0.02

Note: Forms with a sample size ≥ 50 were included in the calculation.

Table 6 shows that in 2020, the mean reliability estimate for the Listening forms was 0.80 and the mean reliability estimate for the Reading forms was 0.82. These values suggest high internal consistency for the CAEL Listening and Reading test forms. The mean distance from the median measures the variability of reliability among test forms. The results indicate that, for both the Listening and Reading components, the forms used in 2020 perform similarly in terms of their reliability. Overall, these values suggest good internal consistency of measurement both within and across CAEL Listening and Reading forms.

3.2.2 Rater Agreement for the Writing and Listening Components

The raters for the Speaking and Writing components of the CAEL Test are highly proficient in English and are fully trained and certified by Paragon. Each speaking and writing task is rated independently by two accredited raters. If the scores awarded by the two raters are in disagreement, the task is evaluated by a third rater. Since the evaluation process for the Writing and Speaking components relies on human judgements and the interpretation and application of a rating scale, variations in judgements are to be expected.

Paragon regularly monitors the reliability of the Speaking and Writing scores of the CAEL Test through rater agreement - the percentage of ratings that are in agreement. Table 8 shows the rater agreement for the Writing and Speaking components in 2020. Overall, the results (90.88% agreement in Speaking and 89.24% agreement in Writing) indicate high consistency of judgement between raters. As mentioned earlier, to ensure fairness and accuracy of all test-taker scores, additional raters will assess the responses when the initial ratings do not agree.

Table 8: Rater Agreement for Speaking and Writing (%)

Component	Agreement
Writing	89.2
Speaking	90.9

4 Closing Remarks

The CAEL Test is one of the best instruments for identifying students with the adequate English skills for academic success. It provides an authentic representation of language use in a Canadian academic context. Paragon is committed to ensuring the comparability of test results between CAEL at a Test Centre and CAEL Online. A continuing line of research and operational monitoring procedures have been established to maintain the comparability of test results and test-taker experience between the two modes of delivery. For additional information on the recent research Paragon conducted on the comparability between CAEL Online and CAEL at a Test Centre, please refer to the report titled A Novel Multimethod Approach to Investigate Whether Tests Delivered at a Test Centre are Concordant with those Delivered Remotely Online: An Investigation of the Concordance of the CAEL.

Paragon offers a range of test preparation materials, both for free and for purchase, to help test takers prepare for the test. For more information about test registration and preparation, visit the CAEL website at www.cael.ca.

Appendix A: CAEL CE Reporting Scale

Band Score	Listening	Reading	Writing	Speaking	Overall
80-90 Expert	Comprehends lectures as well as an experienced academic listener.	Reads academic texts with ease demonstrating comprehension equal to that of experienced academic readers.	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.	Demonstrates a high level of competence, accuracy, and effectiveness in academic/professional settings.
70 Adept	Understands lectures with apparent ease.	Reads and comprehends academic texts with ease.	Readily responds to the demands of the topic and presents information clearly and logically.	Speaks with ease presenting information clearly and logically.	Uses generally accurate language in most settings; some limitations in flexibility are evident.
60 Advanced	Understands information regarding both main ideas and supporting details. May lack some flexibility and miss some information.	Can interpret information with flexibility.	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.	Displays competence in academic or professional settings.
50 High Intermediate	Processes most of the lecture for general ideas but may miss or misinterpret details. Overall understanding is still somewhat restricted.	Reads and comprehends academic texts with ease.	Addresses the topic to a degree but with limited clarity and cohesiveness.	Speaks with some fluency and flexibility but speaks	Exhibits some competence in academic or professional settings; communication may break down in places.
40 Intermediate	Identifies the meaning of some unfamiliar terms but overall understanding is restricted and uneven.	Can interpret information with flexibility.	Makes links among ideas and addresses the topic but writing lacks clarity and cohesiveness.	unevenly; at times there is a natural easy quality to the response but at other times the response breaks down.	Demonstrates some ability to comprehend and articulate complex ideas and arguments typical of academic or professional settings.
30 High Beginner	Makes sense of some sections of lectures by guessing but overall understanding is limited.	Understands main ideas but may misinterpret information. Can identify some relevant details but reads slowly and with greater effort than most academic readers.	Writes something related to the topic but writing is not predictable and language is restricted.	Can speak with some fluency but without flexibility and with noticeable effort.	Expresses basic ideas about familiar topics in routine settings.
10-20 Low Beginner	Takes some meaning from individual words but overall understanding is sketchy and random.	Understands main ideas but is restricted by limited vocabulary and a lack of familiarity with textual conventions.	Uses words randomly but language is very restricted and/or ungrammatical.	Can speak but with false starts, hesitations and some mispronounced words. Provides studied and careful responses.	Communicates with limited ability.

Appendix B: CAEL Scores by Purpose

Test Purpose	Listening		Reading		Writing		Speaking		Overall	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Employment	67.44	13.89	63.33	17.05	61.71	8.88	67.24	8.03	66.22	10.07
For post-secondary admissions	66.00	17.58	61.55	20.59	59.88	11.73	63.92	12.72	64.13	14.05
Other	66.00	16.67	60.12	17.94	58.12	11.40	65.03	10.91	63.39	12.52

Appendix C: CAEL Scores by the Top 8 Declared Nationalities

Country	Listening		Reading		Writing		Speaking		Overall	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Canada	66.63	14.76	62.11	17.75	61.67	9.98	68.10	9.22	65.90	11.17
India	69.12	12.50	64.37	15.80	61.59	7.71	67.26	6.23	66.91	8.65
China	62.38	23.50	60.60	25.55	56.64	15.11	58.04	16.29	60.98	18.57
Philippines	61.79	13.68	56.74	18.87	59.58	8.86	62.00	9.41	61.47	10.81
South Korea	65.95	14.07	63.10	16.06	59.05	6.88	59.52	8.20	62.62	9.83
Pakistan	67.01	15.96	62.69	17.11	62.99	8.88	69.55	6.61	66.87	10.76
Iran	67.88	17.23	66.06	19.60	59.85	9.03	64.24	9.62	65.76	12.29
Nigeria	69.04	11.76	67.88	16.73	65.00	9.60	68.65	7.93	69.04	9.13

Appendix D: CAEL Scores by the Top 8 Declared First Languages

Language	Listening		Reading		Writing		Speaking		Overall	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Panjabi	66.63	13.41	62.27	16.91	60.53	7.81	65.98	7.13	65.07	9.45
English	70.59	14.12	64.80	16.87	64.95	10.37	72.05	8.60	69.37	10.63
Hindi	69.02	12.75	63.49	15.65	61.20	8.32	68.62	5.79	67.03	8.92
Chinese	60.22	22.76	58.48	25.33	55.58	15.17	56.88	16.21	59.24	18.21
Arabic	62.10	13.59	57.54	17.44	58.36	9.27	64.15	9.12	61.49	11.00
Urdu	65.85	14.70	61.48	16.02	61.69	8.25	68.38	7.01	65.49	10.01
Farsi	64.61	15.70	61.63	19.03	59.79	9.74	64.47	9.21	63.97	11.70
Tagalog	62.22	13.20	56.35	18.40	59.37	8.46	62.30	9.56	61.19	10.48